UNIVERSITY OF LONDON THESIS

Degree PhD Year 2006 Name of Author WHINNETT

THE PHYLOGEOGRAPHY AND MOLECULAR EVOLUTION OF ITHOMIINE BUTTERFLIES

by
ALAINE JEAN WHINNETT

A thesis submitted for the degree of
Doctor of Philosophy of the University of London

September 2005

Department of Biology
University College London
4 Stephenson Way
London NW1 2HE

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

This thesis uses molecular techniques to investigate aspects of the evolution of ithomiine butterflies (Lepidoptera: Nymphalidae: Ithomiinae).

1) This thesis takes a comparative phylogeographic approach to investigate the diversification of ithomiines collected across an Amazonian suture zone in N. E. Peru. High variability of mitochondrial DNA (mtDNA) and triosephosphate isomerase (*Tpi*) sequence divergences was recovered, which, i) suggested that diversification of the ithomiines studied here was inconsistent with predictions of the Pleistocene forest refugia theory, one of the leading hypotheses used to explain the record richness of Amazonian biodiversity, and ii) challenged the categorisation of taxa based purely on DNA divergence thresholds, as proposed by DNA barcoding.

2) This thesis also investigates the contribution of ecological adaptation versus allopatric differentiation in explaining the distribution patterns of 4 subspecies belonging to the ithomiine species *Hyposcada anchiala*. The mtDNA sequence data revealed that the most recent radiations were consistent with allopatric divergence during the Pleistocene.

3) In addition, this thesis generates gene genealogies for the ithomiine tribe Oleriini, based on regions of mtDNA, *wingless* and *elongation factor 1*-α. In nearly all cases individuals were clustered by species into the four recognised genera. However, the relationships between the genera remains undetermined. These data contribute to a complete Oleriini phylogeny, which will be used to examine aspects of the evolution of this tribe.

4) Finally, this thesis contributes to the development of nuclear loci for PCR in Lepidoptera. *Tpi* had previously been used for phylogenetics, but here was further developed so that a longer region could be amplified. Primers were also developed for a novel region, *Tektin*, which is shown to have phylogenetic utility at the genus, tribe and subfamily levels.

ACKNOWLEDGEMENTS

I would like to extend thanks to the many people, in many countries, who so generously contributed to the work presented in this thesis.

Special mention goes to my enthusiastic supervisor, Jim Mallet. My PhD has been an amazing experience and I thank Jim wholeheartedly, not only for his tremendous academic support, but also for giving me so many wonderful opportunities. Not many PhDs involve a helicopter ride with the Ecuadorian Army into the rainforest, or travelling through Mongolia to get to a conference in Siberia!

Similar, profound gratitude goes to Andy Brower, who has been a truly dedicated mentor. I am particularly indebted to Andy for his constant faith in my lab work, and for his support when so generously hosting me in Oregon. I have very fond memories of my time there.

I am also hugely appreciative to Keith Willmott, especially for sharing his taxonomic expertise so willingly, and for being so dedicated to his role as my secondary supervisor.

Special mention goes to Marga Beltrán, Russ Naisbit, Marie Zimmermann, Vanessa Bull, Ming-Min Lee, Ian Evans, Jacques Gianino, Chris Jiggins, Gerardo Lamas and Fraser Simpson, for going far beyond the call of duty. To Nick Mundy, for encouraging me to embark on the molecular biology path, and for providing me with a fantastic lab training. And to Vernon Reynolds, Peter Davies and Mr Sumner, for nurturing my enthusiasm for biology.

Finally, but by no means least, thanks go to mum, dad and Steve for almost unbelievable support. They are the most important people in my world and I dedicate this thesis to them.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Chapter Six

Appendix One

Appendix Two

LIST OF TABLES

Chapter Two

Chapter Three

Chapter Five

Chapter Six

Appendix One

Appendix Two

DECLARATION

Chapters 1, 3, 6 and 7

The work presented in Chapters 1, 3, 6 and 7 is entirely my own.

Chapter 2

A version of the work presented in Chapter 2, entitled 'Strikingly variable divergence times inferred across an Amazonian butterfly 'suture zone'' has now been published, in *Proceedings of the Royal Society of London, Series B* (2005, **272**: 2525-2533). As of July 2006, this article does not have any citations. I am the principle author of the paper, which is co-authored with (in order on the publication): Marie Zimmermann (University College London (UCL), London, UK); Keith Willmott (Natural History Museum, London, UK); Nimia Herrera and Ricardo Mallarino (Smithsonian Tropical Research Institute, Balboa, Panama); Fraser Simpson (UCL, London, UK): Mathieu Joron (University of Edinburgh, Edinburgh, UK); Gerardo Lamas (Universidad Nacional Mayor de Museo de Historia Natural San Marcos, Lima, Peru); and Jim Mallet (UCL, London, UK).

I headed preparations for the field work trip to Peru in 2002, which included successfully obtaining a grant to fund a field assistant, and obtaining collection plus export permits. I co-headed the first 45 field days with Marie and Jim, and headed the final 30 field days. I also co-ordinated the identification and cataloguing of the 2500+ preserved specimens. From this collection, I selected the 98 individuals from the ithomiine genera: *Dircenna*, *Hyalyris*, *Brevioleria*, *Hyposcada*, *Ithomia*, *Napeogenes*, *Oleria* and *Pseudoscada*, to be included in the study. I made an application to, and was awarded a fellowship at The Smithsonian Tropical Research Institute, to cover my laboratory expenses. There, I obtained mtDNA sequence data for 90 of the 98 selected individuals (8 *Ithomia* sequences were already available from GenBank, see below). I lodged 164 sequences in Genbank (DQ078312- 473; DQ078478- 479) (the 90 I had generated, plus the 74 contributed by Marie, as detailed below). I then compiled a final data set of these sequences, plus the 8 *Ithomia* sequences and the outgroup. I ran all initial analyses, including testing for constancy of evolutionary rates and generating the Bayesian tree (Figure 2). I interpreted the findings and wrote the manuscript.

The co-authors contributed as follows: Jim initiated the project and sourced most of the funding for the field work. Marie, Jim, Fraser, and Mathieu collected

specimens during the trip in 2002. Marie and Gerardo identified all the *Melinaea* and *Hypothyris*. Gerardo, Keith and Jim coached me in the identification, and provided information about hybridisation, of all ithomiines in the San Martin region. They also checked all the identifications that I had made, to subspecies level, and identified all of the specimens that I was unable to recognise. Fraser entered the collections into a database upon return to the UK. Marie selected and sequenced the majority of the 42 *Melinaea* and 32 *Hypothyris* specimens included in the study. I aided Marie to do this, for example, by optimising the PCR parameters, and by working on those individuals which proved particularly difficult to amplify. Nimia contributed by running the sequencing gels. Ricardo provided me with lab protocols, as well as sequence data prior to its publication for 8 specimens of *Ithomia* (AY713067- 8, AY713040- 1, AY713075- 8). Jim, Marie and Keith contributed much to the discussions regarding the context of the study, including Amazonian diversification and DNA 'barcoding'. They also provided formative feedback on the original manuscript. Jim lead the rate-smoothed analyses, which were performed in accordance with the recommendations from peer-reviewers. All co-authors proof read the article prior to its final submission.

Chapter 4

A version of Chapter 4, entitled 'The phylogenetic utility of *Tektin*, a novel region for inferring systematic relationships amongst Lepidoptera', has been now published, in *Annals of the Entomological Society of America* (2005, **98**: 873-886). As of July 2006, this article does not have any citations, although it was referred to whilst in press by: Mallarino, R., Bermingham, E., Willmott, K. R., Whinnett, A. & Jiggins, C. D. (2005) Molecular systematics of the butterfly genus *Ithomia* (Lepidoptera: Ithomiinae): a composite phylogenetic hypothesis based on seven genes. *Molecular Phylogenetics and Evolution*, **34**: 625-644.

I am first author of the paper, which is co-authored with: Andrew Brower and Ming-Min Lee (Oregon State University, Corvallis, USA); Keith Willmott (Natural History Museum, London, UK); and Jim Mallet (University College London, London, UK). The central tenet of this chapter is the work on *Tektin*, for which I carried out all of the *Tektin* primer development, generated all of the *Tektin* sequences and lodged these in GenBank (AY848711- 46). I also performed all of the *Tektin* and comparative data analyses. Many of the comparative sequences studied here were already available, the authorships of which are detailed in Table 1 of the chapter. The contributions made

by the co-authors are as follows: whilst hosting me in his laboratory, Andrew provided financial support for the *Tektin* primer development as well as *Tektin* sequencing, and granted me permission to use template DNA from his collection. Andrew also contributed greatly by providing me with access to comparative data, for wingless, elongation factor 1-α and mitochondrial regions, which had already been generated but were prior to publication. Of these sequences, many have now been published, in: Brower, A. V. Z., Freitas, A. V. L., Lee, M.-M., Silva Brandão, K. L., Whinnett, A. & Willmott K. R. (2006) Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions. *Systematic Entomology*, 31: 288-301. Of the sequences provided by Andrew, 10 were not included in the above mentioned Systematic Entomology publication, but were directly submitted to GenBank by Andrew so that they could be included in the *Tektin* publication. Ming-Min contributed to this chapter by providing technical assistance in the laboratory, such as preparing stock reagents. She also contributed by generating much of the sequence data provided by Andrew. Keith contributed by identifying specimens and providing advice regarding which taxa to include. In particular, Keith provided valuable information about phylogenetic expectations based on morphological characters. Both Keith and Jim Mallet thoroughly reviewed the chapter prior to its journal submission.

Chapter 5

A version of the work presented in Chapter 5, entitled 'Mitochondrial DNA provides an insight into the mechanisms driving diversification in the ithomiine butterfly *Hyposcada anchiala* (Lepidoptera : Nymphalidae : Ithomiinae)' has now been published in the *European Journal of Entomology*, (2005, **102**: 633-639). As of July 2006, this article does not have any citations. I am the principle author of the paper, which is co-authored with (in order on the publication): Keith Willmott (Natural History Museum, London, UK); Andrew Brower (Oregon State University, Corvallis, Oregon); Fraser Simpson and Marie Zimmermann (University College London (UCL), London, UK); Gerardo Lamas (Universidad Nacional Mayor de Museo de Historia Natural San Marcos, Lima, Peru); and Jim Mallet (UCL, London, UK).

I designed the study, selected and obtained all the samples, and performed the lab work on all specimens except for the individual *Hyposcada anchiala richardsi* (G1). I lodged all the sequences in Genbank (DQ078312, DQ078355-62, DQ078474-77), compiled the data set and ran all analyses. I interpreted the findings and wrote the manuscript.

The co-authors contributed as follows: Fraser, Marie, and Jim helped me to collect specimens during field work in 2002, and Gerardo donated individual G1. Gerardo and Keith identified, or checked my identification of, the specimens. Andrew provided financial support for the lab work, and Fraser performed the PCR and cycle sequencing reactions for individual G1. Keith contributed much to discussions about butterfly species richness, provided information about the geographical distributions of *Hyposcada anchiala* subspecies and designed Figure 1. Keith, Jim and Andrew provided formative feedback on the original manuscript, and all other co-authors proof read the article prior to its final submission.

Appendix 1

I am the fourth author of the publication entitled 'Molecular systematics of the butterfly genus *Ithomia* (Lepidoptera: Ithomiinae): a composite phylogenetic hypothesis based on seven genes', submitted here as Appendix 1. I contributed by providing specimens of *Ithomia* which I had collected, by providing the *Tektin* primers prior to their publication, by contributing to discussions and by reviewing the manuscript prior to its submission. Ricardo Mallarino (Smithsonian Tropical Research Institute, Balboa, Panama) performed all of the PCR, sequencing, and much of the analyses, he is the first author of this publication. Other co-authors are: Eldredge Bermingham (Smithsonian Tropical Research Institute, Balboa, Panama), Keith R. Willmott (Natural History Museum, London, UK), and Chris D. Jiggins (University of Edinburgh, Edinburgh, UK).

Appendix 1 has been published in *Molecular Phylogenetics and Evolution* (2005, 34: 625-644). As of July 2006, it has been cited in 5 publications:

(1) Brower, A. V. Z., Freitas, A. V. L., Lee, M.-M., Silva Brandão, K. L., Whinnett, A. & Willmott, K. R. (2006) Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions. *Systematic Entomology*, **31**: 288-301.

(2) Du, Y. L., Roe, A. D. & Sperling, F. A. H. (2005) Phylogenetic framework for Dioryctria (Lepidoptera : Pyralidae : Phycitinae) based on combined analysis of mitochondrial DNA and morphology. *Canadian Entomologist*, **137**: 685-711.

(3) Whinnett, A., Willmott, K. R., Brower, A. V. Z., Simpson, F., Zimmermann, M., Lamas, G. & Mallet, J. (2005) Mitochondrial DNA provides an insight into the mechanisms driving diversification in the ithomiine butterfly Hyposcada anchiala (Lepidoptera : Nymphalidae : Ithomiinae). *European Journal of Entomology*, 102: 633-639.

(4) Whinnett, A., Zimmermann, M., Willmott, K. R., Herrera, N., Mallarino, R., Simpson, F., Joron, M., Lamas, G. & Mallet, J. (2005b) Strikingly variable divergence times inferred across an Amazonian butterfly 'suture zone'. *Proceedings of the Royal Society, Series B*, **272**: 2525-2533.

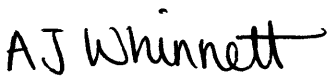
(5) Whinnett, A., Brower, A. V. Z., Lee, M.-M., Willmott, K. R. & Mallet, J. (2005a) The phylogenetic utility of *Tektin*, a novel region for inferring systematic relationships amongst Lepidoptera. *Annals of the Entomological Society of America*, **98**: 873-886.

Appendix 2

I am the fifth author of the publication entitled 'Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions' submitted here as Appendix 2, and published in Systematic Entomology (2005, **31**: 288-301). As of July 2006, this article has no citations.

This publication is senior authored by Andrew Brower (Oregon State University, Corvallis, USA), who led the research programme. Additional co-authors are as follows: André Freitas (UNICAMP, Campinas, Brazil); Ming-Min Lee (Oregon State University, Corvallis, USA); Karina Silva Brandão (UNICAMP, Campinas, Brazil), and Keith Willmott (Natural History Museum, London, UK). I contributed by: providing preserved specimens which I had collected; performing DNA extractions; and generating elongation factor 1-α sequence data for 35, and wingless sequence data for 40, individuals of ithomiine butterfly. I also contributed to discussions about lab techniques, and reviewed the manuscript prior to its publication.

Signed:



Alaine Whinnett
Candidate

Professor James Mallet
Supervisor

CHAPTER ONE

INTRODUCTION

Ithomiine butterflies

Ithomiine butterflies belong to a diverse, Neotropical clade which includes approximately 355 species (Lamas 2004). This grouping is defined by a distinctive synapomorphy of costal androniał hairs on the dorsal hindwing surface of males (Fox 1940). The monophyly of this group has since been supported by additional morphological characters (Freitas & Brown 2004) and molecular data (Brower 2000). Ithomiines have been considered as a tribe in the Danainae subfamily (Ackery et al. 1999, Brower 2000, Brower et al. 2006). However, this thesis follows the most recent checklist for this group (Lamas 2004) and therefore considers ithomiines as belonging to the subfamily Ithomiinae, of the family Nymphalidae (Ehrlich & Ehrlich 1967, Lamas 2004). Within subfamily classification has been complex, largely due to convergent evolution and geographical polymorphisms, and Ithomiinae taxonomy has been reconsidered (reviewed in Brower et al. 2006).

Ithomiines have caught the attention of lepidopterists and other natural historians for many years, and have long been prized specimens in serious entomology collections. Current interest in the Ithomiinae can be exemplified by it being one of just two subfamilies represented on the on-line butterfly database at The Smithsonian National Museum of Natural History (http://entomology.si.edu/entomology/Butterflies/search.lasso).

Ever since the mid-1800s, Ithomiines have been a useful paradigm for a number of studies on evolution and speciation, including the important pioneering work on the theory of mimicry (Bates 1862, Müller 1879), which is considered an important test of the theory of natural selection (Darwin 1863, Wickler 1968). Research into ithomiine butterflies has also addressed the pyrrolizidine alkaloids acquired from plant sources, which are used as pheromone precursors and to chemically protect ithomiines from predators. In addition, ithomiines have been used as key taxa in the debate about the mechanism, or mechanisms, driving Neotropical diversification.

Ithomiine butterflies and mimicry

Based on observations that Amazonian butterflies in particular localities often share colour patterns although they are distantly related, and also that colour patterns differed remarkably over geographic area, Bates developed a theory to explain mimicry (1862). This theory, now known as Batesian mimicry, explains how a palatable species (the mimic) can gain protection from predators which have learnt to avoid a signal advertising another species' (the model's) unpalatability, by converging on that warning pattern. A few years later, Müller (1879) presented a mathematical argument to explain why convergent evolution might arise between species although they are both already protected from predators by distastefulness. His theory, now known as Müllerian mimicry, predicts that such species (co-mimics) gain enhanced protection by sharing a warning signal as they split the loss of those individuals which are attacked before the predator learns to avoid the signal.

Pyrrolizidine alkaloids acquired from plant sources have been experimentally shown to protect adult ithomiines from birds and a spider predator, adding to the evidence that ithomiine butterflies are aposematically (warningly) coloured. Ithomiines often dominate mimicry 'rings' (groups of species that have converged on a colour pattern), as both Batesian models and Müllerian co-models and co-mimics (Bates 1862, Müller 1879, Beccaloni 1997a).

Ithomiines from multiple, diverse mimicry rings are often found in a single location. Beccaloni (1997b) found the 56 ithomiine species at the Jatun Sacha Biological Station (Napo, Ecuador) belonged to eight discrete mimicry complexes, namely the: "clearwing", "orange-tip", "small and dark transparent", "small and yellow transparent", "large and yellow transparent", "yellow-bar tiger", "orange and black tiger", and "tiger" complexes. Ithomiine species can be polymorphic for morphs which participate in different mimicry rings (as in seven species at Jatun Sacha), and that multiple Müllerian mimicry rings co-exist, are both strongly at odds with a simple prediction of mimicry theory, namely that all Müllerian mimics will ultimately converge on a single warning signal. This expectation arises from Müllerian mimicry having greatest value when the frequency of individuals displaying a given warning signal is highest, as the

warning signal is more strongly reinforced and the attack rate per individual is at its lowest (Mallet & Barton 1989, Mallet & Joron 1999).

However, this classic model of mimicry oversimplifies situations in nature such as habitat heterogeneity, interacting taxa, and temporal fluctuations. There is mounting evidence that apparently co-existing mimicry rings are never completely sympatric, but that they are separated into distinct microhabitats. For example, different ithomiine mimicry rings have been shown to be numerically dominant at distinct 'height bands' and further segregated by vegetation type (Papageorgis 1975, Mallet & Gilbert 1995, Medina et al. 1996, Beccaloni 1997a, DeVries et al. 1999, Willmott & Mallet 2004). Secondly, ithomiine butterflies are not evolving in isolation; also in the eight mimicry complexes at Jatun Sacha are the 69 species of ithomiine mimics (34 butterfly, 34 moth and one damselfly species). The so-called 'escape hypothesis' predicts that ithomiines might gain an advantage by breaking away from a signal shared with Batesian or quasi-Batesian (less unpalatable Müllerian) mimics as patterns shared with these mimics offer a reduced protection, due to predators learning the warning signal more slowly if first encountering the more palatable mimics (Pough et al. 1973, Beccaloni 1997b). In addition, the mimetic environment fluctuates, so that particular patterns might have a higher protection at different times. Thus, co-mimics might benefit from switching to the current, locally most highly protected mimicry ring. This is supported by long-term changes in abundance, and even presence, of ithomiines and their co-mimics over time and space (Brown & Benson 1974, Joron et al. 1999). This subject, as well as other aspects of the evolution of mimicry in butterflies, is addressed in more detail by Joron and Mallet (1998) and Mallet and Joron (1999).

Mimicry has been implicated in driving the diversification of butterflies, for example in cases where there is strong mating preference for colour (Mallet et al. 1998, Jiggins et al. 2001). Initial shifts in colour pattern can be re-enforced by the non-mimetic hybrids which are subject to strong purifying selection (Naisbit et al. 2003). Jiggins and colleagues (2006) addressed the role that mimicry plays in driving speciation of the ithomiine genus *Ithomia* from a phylogenetic perspective. They identified just one speciation event as a candidate for mimicry-driven diversification. In addition, some clades were identified where speciation had occurred without a colour pattern switch, and several

polymorphic species of *Ithomia* were found to share at least one of their colour patterns with their closest relative, suggesting those colour patterns not shared with the closest relative arose after speciation. Based on the above, Jiggins and colleagues (2006) suggested that although colour pattern change can cause speciation, speciation in *Ithomia* can occur without colour pattern change and therefore 'there seems little evidence that recent speciation events have involved switches in colour pattern [in *Ithomia*]'. Interestingly, those branches where colour pattern change occurred were significantly shorter than branches without a pattern change, suggesting that when colour pattern is involved, it might promote more rapid and repeated speciation than other mechanisms (Jiggins et al. 2006).

Although efforts have not yet been made to specifically investigate the genetic basis of mimicry in ithomiines, exciting research has addressed this in *Heliconius*. *Heliconius* belong to the subfamily Heliconiinae, and like the Ithomiinae are in the family Nymphalidae. Based on breeding experiments, largely on the co-mimetic *Heliconius erato, Heliconius cydno* and *Heliconius melpomene*, a small number of dominant or semidominant loci are reported to have a major effect on the colour patterns, the sizes and shapes of which are modulated by minor genes. It is thought that similar colour patterns in these species sometimes arise through shared mechanisms, but that there are also instances of phenotypic similarity resulting from different genetic mechanisms (Sheppard et al. 1985, Mallet 1989, Mallet 1993, Jiggins & McMillan 1997, Gilbert 2003, Naisbit et al. 2003). In contrast, crosses in *H. numata* have indicated that all colour polymorphisms in this species are determined by just a single gene or supergene. This locus has a simple dominance hierarchy for its multiple alleles which prevents the production of non-mimetic intermediates (Joron et al. 1999). Genetic linkage maps have recently been described for *H. erato* (Tobler et al. 2005) and *H. melpomene* (Jiggins et al. 2005), and these are currently being used comparatively to characterise the genes controlling colour pattern diversity. The colour pattern gene in *H. numata* has also been located, interestingly to the same region as a colour pattern gene already known in *H. melpomene* (Joron, pers. comm.). Research into the ecological and evolutionary genomics of *Heliconius* can be followed on http://heliconius.org/.

Ithomiine butterflies and pyrrolizidine alkaloids

Pyrrolizidine alkaloids (PAs) are thought to be used by male ithomiines as pheromone precursors (see Edgar et al. 1976, Boppre 1986, Eisner & Meinwald 1987, Trigo et al. 1994). In support of this, Schulz and colleagues (2004) found 13 volatile compounds formed from the PA lycopsamine, in the pheromone releasing male hairpencils from 30 ithomiine genera. Ithomiine pheromones are important in attracting both male and female ithomiines, and initiating ithomiine lek formation (Haber 1978).

The majority of research into ithomiines and PAs has addressed the role of PAs in chemical defence. "Chemical defence can be suggested when individual prey organisms contain one or more noxious chemical substances which facilitate proximal and/or distal rejection by predators; rejection can occur after a predator partially to completely ingests one or more prey individuals, or after the predator simply smells or tastes the prey" (Brower 1984). PAs are considered to be the primary chemical defence mechanism in ithomiine butterflies (Brown 1987). *Tithorea harmonia* and *Aeria olena* are the only ithomiine species known to incorporate PAs from larval foodplants. Most ithomiine species sequester PAs as adults, from the exudates and nectar of plants belonging to the Asteraceae and Boraginaceae (Pliske 1975, Brower 1984, Brown 1985, Boppre 1986, Brown 1987, Trigo et al. 1996). In general, male ithomiines are highly attracted to these PA sources, and sequester them directly. Females are less attracted to PA sources, although they probably collect some PAs from plant sources (Masters 1990). Females are thought to gain protective PAs from males, via the spermatophore received during mating. The eggs are protected as a result of females transmitting PAs to the egg shell (Boggs & Gilbert 1979, Brown 1987).

PAs have been experimentally shown to be unpalatable to birds, including the pileated finch *Coryphospingus pileatus* (Cardoso 1997), and also to the orb-weaving *Nephila clavipes* spider (Silva & Trigo 2002). Orb-weaving spiders eat palatable butterflies; Heliconiini, Acraeinae, Nymphalinae, Pieridae, Papilonidae, and even freshly emerged ithomiine adults. However, these same spiders have been observed to rapidly cut older ithomiine adult butterflies (that had sequestered PAs) out of their webs after any contact with the body or wings, and then release them unharmed (Vasconcellos-Neto & Lewinsohn 1984, Brown

1985, Brown 1987). Observations that freshly emerged adults of *Tithorea harmonia* were eaten (this species is atypical as it sequesters small amounts of PAs from larval foodplants), but that field caught adults were released by the orb-weaving spider, led Trigo and colleagues (1996) to conclude that protection against predation by the orb-weaving spider may be partly dependent on concentration of PAs in the butterfly. Active investigations into ithomiines and PAs continue, with researchers at Universidade Estadual de Campinas, Brazil currently leading in much of this work.

Ithomiine butterflies and Neotropical diversification theories

The diversification of taxa in temperate zones is quite well understood, with temperate glacial refuges well documented and widely accepted (Hewitt 2000). However, there is little agreement on which factors promoted species diversification in the Neotropics. A number of theories have been proposed to explain Neotropical diversification, below I summarise the leading hypotheses; vicariant and ecological. Although focussing on particular mechanisms may be overly simplistic (Haffer 1997, Cheviron et al. 2005), elucidating which particular theory or theories contribute most significantly to the generation of Neotropical biodiversity is critical for advancing our understanding of the evolution in this region.

Ithomiines have played an important role in this debate, as a number of authors have postulated that ithomiine distribution reflects allopatric speciation of previously widespread taxa. This is consistent with, and has been used to support, the Pleistocene refugia vicariance model of diversification (Turner 1971, Brown 1979, Brown 1982, Sheppard et al. 1985, Brown 1987, Turner & Mallet 1996).

Vicariance hypotheses

Vicariance theories hold that that evolutionary separation is determined by geographical separation. Physical geographical vicariance theories concern the formation of features such as mountain ranges, valleys or expanses of water. These split previously widespread taxa, and form effective barriers to dispersal so that taxa on either side of the geographical formation evolve independently (Wiley 1988).

It has long been recognised that Amazonian rivers can act as a barrier to taxa: 'Rivers generally do not determine the distribution of species, because, when small, there are few animals which cannot pass them but in very large rivers the case is different, and they will, it is believed, be found to be the limits, determining the range of many animals of all orders. With regard to the Amazon, and its larger tributaries, I have ascertained this to be the case' (Wallace 1853). Such biogeographical observations led to the development of the Riverine barrier (or river) hypothesis as an explanation for the record richness of Amazonian biodiversity. The Riverine barrier hypothesis contends that previously widespread ancestral Amazonian taxa were separated when large Amazonian rivers formed during the Late Tertiary and Early Quaternary. The taxa on each side of the dividing river were hypothesised to then develop in isolation (Sick 1967, Capparella 1988).

Predictions of the Riverine barrier hypothesis, that sister lineages occur across major rivers, and that wider stretches of the river provide a stronger barrier, are amenable to testing with molecular data. Based on studies of saddle-back tamarins (*Saguinus fuscicollis*) which inhabit forest on both sides of the Rio Jurua, Peres and colleagues (1996) reported that gene flow occurred between tamarins across the river. However, this gene flow was restricted to the headwater section of the river, and is therefore consistent with expectations of the Riverine barrier hypothesis. In another study on the Spix's (elegant) woodcreeper superspecies (*Xiphorhynchus spixii/elegans*), Aleixo (2004) reported that sister lineages were separated by rivers on the Brazilian shield, as well as in central and eastern Amazonia, also consistent with expectations of the Riverine barrier hypothesis. However, this finding did not apply to woodcreepers located in western Amazonia, where predictions of other hypotheses were more strongly supported.

There are a number of problems with the strict Riverine barrier hypothesis, including; that land and its inhabitants can be transferred across rivers, for example, when meander loops are cut off, or new river courses are cut; and that some taxa which are separated by rivers, can traverse other larger rivers, indicating that rivers are not *per se* barriers for those taxa (Haffer 1997). These issues along with a number of genetic studies which are largely inconsistent with the Riverine hypothesis, based on; small bodied mammals and frogs (Gascon et

al. 2000); small bodied, non-flying mammals (Patton et al. 2000); the dart-poison frog *Epipedobates femoralis* (Lougheed et al. 1999); and the arboreal spiny rat, *Mesomys hispidus* (Patton et al. 1994); have resulted in the Riverine barrier hypothesis as an explanation for Neotropical diversity, except for primates, being largely discounted today (Haffer 1997).

The dominant biogeographical vicariance theory is the refuge theory. This theory suggests that dry periods would have forced previously widespread forests into fragments isolated by open vegetation. The taxa diverged within these forest fragments so that when the refuges re-expanded and brought the taxa from the different fragments back into contact, they were partially or completely reproductively isolated.

The Pleistocene refuge hypothesis contends that dry conditions specifically in the Pleistocene led to allopatric speciation in forest refuges (Haffer 1969). Although it is frequently used to explain diversification in the Amazon basin, only a few molecular studies have implicated Pleistocene refuges as responsible for the current distribution patterns, and possibly in the historical isolation, of Amazonian taxa. One example is the recent study by Wuster and colleagues (2005) on the Neotropical rattlesnake, *Crotalus durissus*.

Despite identifying genetic evidence of the population growth associated with habitat expansions after the late Pleistocene in North American mammals, Lessa and colleagues (2003) were unable to do the same for small Amazonian mammals. The Pleistocene refuge theory has also failed to account for diversification in other taxa, for example, *Pionopsitta* parrots, *Pteroglossus* toucans (Eberhard & Bermingham 2005), and *Heliconius* butterflies (Flanagan et al. 2004). Such reports are in accordance with palaeoecological data from the Amazon basin (Colinvaux 1997, Colinvaux & De Oliveira 2000, Colinvaux et al. 2000, Colinvaux & De Oliveira 2001) which suggest that the climatic oscillations of the Pleistocene did not have a dramatic effect on forest cover. In addition, the Pleistocene refuge hypothesis is under attack from the increasing number of molecular studies which suggest that taxa originated prior to the Pleistocene (Collins & Dubach 2000, Glor et al. 2001, Lovette 2004, Eberhard & Bermingham 2005).

Ecological or gradient hypotheses

The Neotropical rainforest is very ecologically diverse and contains steep environmental gradients. Proponents of ecological hypotheses argue that environmental variations affect taxa in different ways and that this can drive parapatric diversification of taxa (Endler 1982). Although a number of researchers have reported their data to be inconsistent with ecological explanations (Prum 1988, Cracraft & Prum 1988, Patton & Smith 1992), others support ecological explanations, for example, Svenning (2001) who reported heterogeneous gradients as 'an important diversity-generating factor in the Neotropical palm flora'.

In their recent book on speciation, Coyne and Orr (2004) suggested that sister species with contiguous ranges that meet at an ecotone, and were never separated by a geographic barrier, can be better explained by parapatric speciation (consistent with ecological hypotheses) than allopatric speciation (as in vicariance hypotheses). They stated "The evidence becomes stronger if this pattern is obeyed by sister species in several diverse groups, and if divergence times between members of different sister pairs are not congruent, suggesting that their ancestors were not divided by a geographic barrier at the same time. Such a conclusion will require accurate phylogenies and good estimates of divergence times, and we know of no study of parapatric speciation that satisfies these conditions." In the following chapter, I attempt to undertake such a study.

Scientific rationale and thesis outline

Swift progress has been made in molecular research and some even consider current times to be 'the age of molecular biology'. Despite intense interest in ithomiines, at the onset of this thesis research into these butterflies at the molecular level was limited to a single assessment of the genetic diversity of two subspecies of *Oleria onega* using the method random amplification of polymorphic DNA (RAPD) (since published, Gallusser et al. 2004). As already discussed, diverse aspects of ithomiine biology have been investigated over more than a century, making the Ithomiinae one of the best understood neotropical groups to date. Given that molecular research into the Ithomiinae had been so limited, and that a wealth of information from other research areas was available to maximise the significance of any molecular findings, it was of considerable interest to investigate ithomiine butterflies from a molecular perspective.

Here, the chapters and appendices are briefly introduced. In Chapter 2, mitochondrial sequence data for *cytochrome oxidases I* and *II*, plus the intervening tRNA (COI-tRNA-COII) were used to estimate divergence times between ithomiines sampled from N. E. Peru. The inferred divergence times differed remarkably at each taxonomic level, a finding in striking contrast to the correlated lineage splits predicted by the Pleistocene forest refugia theory. Therefore, the studied ithomiines did not appear to conform to one of the leading hypotheses used to explain the richness of Amazonian biodiversity. The high variability of sequence divergences also challenges the categorisation of taxa based purely on DNA divergence thresholds, as proposed by systematists who advocate the use of DNA barcodes.

In Chapter 3, the *triosephosphate isomerase* (*Tpi*) locus, which had previously been proven to be phylogenetically useful (using primers based in exons 3 and 4), was further developed to enable a longer region to be amplified (exon 1 through 5). *Tpi* was then used to investigate if the findings from Chapter 2 would be corroborated by this independent locus. A marked variability in pairwise divergences for comparisons at different taxonomic levels was recovered, mirroring the previous findings based on mitochondrial DNA. Features of the molecular evolution of the *Tpi* and COI-tRNA-COII regions were also compared.

Primers were developed for an 807 base pair, coding nuclear region homologous to the testis-specific *Tektin* gene in *Bombyx mori* in Chapter 4. The *Tektin* primers were used on 34 taxa from the Ithomiinae, Danainae and Heliconiinae, and were shown to have phylogenetic utility at the genus, tribe and subfamily levels. The *Tektin* topology was in broad agreement to traditional classification and a recent cladogram inferred from morphology. In addition, the *Tektin*, COI-tRNA-COII, *wg* and *elongation factor 1-α* (*Ef1α*) regions were compared, to assess their relative utilities in inferring relationships at different phylogenetic levels.

Phylogenies representing five subspecies of *Hyposcada anchiala* were created using a 1567 bp alignment of COI-tRNA-COII in Chapter 5. These were used to investigate the possible roles of ecological adaptation versus allopatric differentiation in explaining a distributional pattern of two submontane subspecies, both of which have a melanic-tiger wing patterning (*Hyposcada anchiala mendax* in N. Peru and *Hyposcada anchiala fallax* in S. Peru), and two lowland subspecies both of which have a tiger wing patterning (*Hyposcada anchiala interrupta* in N. Peru and *Hyposcada anchiala richardsi* in S. Peru). The sequence data revealed that the most recent radiations were consistent with allopatric divergence dated to the Pleistocene.

In Chapter 6, gene genealogies are presented for 41 of the approx. 60 species belonging to the tribe Oleriini, based on data from *Ef1α*, *wingless* (*wg*), and the mitochondrial COI-tRNA-COII. In nearly all cases, individuals clustered by species into the four recognised genera: *Hyposcada*, *Megoleria*, *Ollantaya* and *Oleria*. However, the relationships between these genera remain unclear. The *Oleria* had previously been assigned into species groups based on morphological characters. The phylogenetic hypotheses presented here support one of the species groups, but additional molecular data are required to determine if the other groups are also supported.

In Appendix 1, species-level phylogenetic hypotheses are presented for seven gene regions, including the *Tektin* locus developed in Chapter 4, for the genus *Ithomia*. The combined evidence topology had a strong general relationship with geography, for example, the clade containing *Ithomia agnosia* was largely restricted to southeastern Brazil and the Amazon basin. The inferred

relationships have since been used to address modes of speciation and to reconstruct the history of wing colour pattern evolution within this genus (Jiggins et al. 2006). It should be noted that this paper considers the mitochondrial COI-tRNA-COII as three distinct regions, even though the mitochondrial genome usually evolves as one unit, and COI, tRNA and COII are combined for analysis. This contiguous stretch of the mitochondrial genome is considered as one region in the rest of this thesis.

In Appendix 2, phylogenetic hypotheses based on COI-tRNA-COII, *wg* and *Ef1 α*, are presented for 81 species of 41 ithomiine genera, and ithomiine relationships are assessed. Overall, the topology derived from molecular data was very similar to previously presented morphological topologies. However, a number of important differences provoke a reconsideration of some prior hypotheses. For example, the Apocynaceae feeding Tithoreina were previously considered as sister to the other ithomiines, which are Solanaceae feeders (it should be noted that although ithomiines are regarded to be a subfamily throughout the rest of thesis, they are considered to be a tribe in appendix 2). However, molecular data revealed the Melinaeina to be the sister group to all other ithomiines, implying that Apocynaceae feeding in the Tithoreina was an independent shift back to the host plant family of the Danainae. The sequence data generated here have also contributed to resolving the relationship among Ithomiinae, Danainae and Tellervini (Brower et al. pers. comm.).

References

Ackery, P. R., de Jong, R. & Vane-Wright, R. I. (1999) The butterflies: Hedyloidea, Hesperioidea and Papilionoidea. In (Ed.) N. P. Kristensen, Lepidoptera, moths and butterflies. 1. Evolution, systematics and biogeography. *Handbook of Zoology,* **4** (35): Lepidoptera, pp. 263-300. de Gruyter, Berlin.

Aleixo, A. (2004) Historical diversification of a Terra-firme forest bird superspecies: A phylogeographic perspective on the role of different hypotheses of Amazonian diversification. *Evolution,* **58** (6): 1303-1317.

Bates, H. W. (1862) Contributions to an insect fauna of the Amazon Valley. Lepidoptera: Heliconidae. *Transactions of the Linnean Society,* **23**: 495-566.

Beccaloni, G. W. (1997a) Vertical stratification of ithomiine butterfly (Nymphalidae: Ithomiinae) mimicry complexes: the relationship between adult flight height and larval host plant height. *Biological Journal of the Linnean Society,* **62**: 313-341.

Beccaloni, G. W. (1997b) Ecology, natural history, and behaviour of ithomiine butterflies and their mimics in Ecuador (Lepidoptera: Nymphalidae: Ithomiinae). *Tropical Lepidoptera,* **8**: 103-124.

Boggs, C. L. & Gilbert, L. E. (1979) Male contribution to egg production in butterflies: evidence for transfer of nutrients at mating. *Science,* **206**: 83-84.

Boppre, M. (1986) Insects pharmacophagously utilizing defensive plant chemicals (pyrrolizidine alkaloids). *Naturwissenschaften* (Berlin), **73**: 17-26.

Brower, A. V. Z. (2000) Phylogenetic relationships among the Nymphalidae (Lepidoptera) inferred from partial sequences of the wingless gene. *Proceedings of the Royal Society, London B,* **267**: 1201-1211.

Brower, A. V. Z., Freitas, A. V. L., Lee, M.-M., Silva Brandão, K. L., Whinnett, A. & Willmott K. R. (2006) Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions. *Systematic Entomology,* **31**: 288-301.

Brower, L. P. (1984) Chemical defence in butterflies. *The Biology of Butterflies.* (Ed.) P. Ackery and R. I. Vane-Wright, pp. 109-134. Academic Press, London.

Brown, K. S. Jr. (1979) *Ecologia Geográfica e Evolução nas Florestas Neotropicais.* Campinas, São Paulo, Brasil, Universidade Estadual de Campinas.

Brown, K. S. Jr. (1982) Historical and ecological factors in the biogeography of aposematic Neotropical butterflies. *American Zoologist*, **22**: 453-471.

Brown, K. S. Jr. (1985) Chemical ecology of dehydropyrrolizidine alkaloids in adult Ithomiinae (Lepidoptera: Nymphalidae). *Revista Brasileira de Biologia*, **44**: 435-460.

Brown, K. S. Jr. (1987) Chemistry at the Solanaceae/Ithomiinae interface. *Annals of the Missouri Botanical Garden*, **74**: 359-397.

Brown, K. S. Jr. (1987) Areas where humid tropical forest probably persisted. In (Ed.) T. C. Whitmore & G. T. Prance, *Biogeography and Quaternary History in Tropical America*, pp. 45. Oxford University Press, Oxford, U.K.

Brown, K. S. Jr. (1987) Biogeography and evolution of neotropical butterflies In (Ed.) T. C. Whitmore & G. T. Prance, *Biogeography and Quaternary History in Tropical America*, pp. 66-104. Oxford University Press, Oxford, U.K.

Brown, K. S. Jr. & Benson, W. W. (1974) Adaptive polymorphism associated with multiple Müllerian mimicry in Heliconius numata (Lepid.: Nymph.). *Biotropica*, **6**: 205-228.

Brown, K. S. Jr. & Freitas, A. V. L. (1994) Juvenile stages of Ithomiinae: overview and systematics. *Tropical Lepidoptera*, **5**: 9-20.

Capparella, A. (1988) Genetic variation in Neotropical birds, implication for the speciation process. *Acta XIX Congress of International Ornithology (Ottawa 1986)*, **2**: 1658-64.

Cardoso, M. Z. (1997) Testing chemical defence based on pyrrolizidine alkaloids. *Animal Behaviour*, **54** (4): 985-91.

Cheviron, Z. A., Hackett, S. J. & Capparella, A. P. (2005) Complex evolutionary history of a Neotropical lowland forest bird (*Lepidothrix coronata*) and its implications for historical hypotheses of the origin of Neotropical avian diversity. *Molecular Phylogenetics and Evolution*, **36** (2): 338-357.

Collins, A. C. & Dubach, J. M. (2000) Biogeographic and ecological forces responsible for speciation in *Ateles*. International Journal of Primatology, **21** (3): 421-444.

Colinvaux, P. A. (1997) An arid Amazon? *Trends in Ecology and Evolution*, **12**: 318-319.

Colinvaux, P. A. & De Oliveira, P. E. (2000) Palaeoecology and climate of the Amazon basin during the last glacial cycle. *Journal of Quaternary Science*, **15**: 347-356.

Colinvaux, P. A. & De Oliveira, P. E. (2001) Amazon plant diversity and climate through the Cenozoic. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **166**: 51-63.

Colinvaux, P. A., De Oliveira, P. E. & Bush, M. B. (2000) Amazonian and neotropical plant communities on glacial time-scales: The failure of the aridity and refuge hypotheses. *Quaternary Science Reviews,* **19**: 141-169.

Coyne, J. A. & Orr, H. A. (2004) *Speciation.* Sinauer Associates. Sunderland, Massachusetts.

Cracraft, J. & Prum, R. O. (1988) Patterns and processes of diversification: speciation and historical congruence in some Neotropical birds. *Evolution,* **42**: 603-620.

Darwin, C. (1863) A review of H. W. Bates' paper on "mimetic butterflies." In (Ed.) Barrett, P. H. (1977) The Collected Papers of Charles Darwin. Vol. 2. pp. 87-92 University of Chicago Press, Chicago.

DeVries, P. J., Lande, R. & Murray, D. (1999) Associations of co-mimetic ithomiine butterflies on small spatial and temporal scales in a neotropical rainforest. *Biological Journal of the Linnean Society*, **67**: 343-364.

Eberhard, J. R. & Bermingham, E. (2005) Phylogeny and comparative biogeography of *Pionopsitta* parrots and *Pteroglossus* toucans. Molecular Phylogenetics and Evolution, **36**: 288-304.

Edgar, J. A., Cockrum, P. A. & Frahn, J. L. (1976) PAs in *Danaus plexippus* and *Danaus chrysippus* L. *Experientia*, **32**: 1535-1537.

Eisner, T. & Meinwald, J. (1987) Alkaloid-derived pheromones and sexual selection in Lepidoptera. In (Eds.) Prestwich, G. D. & Blomquist, G. J. *Pheromone Biochemistry.* pp. 251-269. Academic Press, New York.

Ehrlich, P. R. & Ehrlich, A. H. (1967) The phenetic relationships of the butterflies. 1. Adult taxonomy and the nonspecificity hypothesis. *Systematic Zoology*, **16**: 301-317.

Endler, J. (1982) Pleistocene forest refuges: fact or fancy? In (Ed.) Prance, G.T. *Biological Diversification in the Tropics.* pp 179-200. New York: Columbia University Press.

Flanagan, N. S., Tobler, A., Davison, A., Pybus, O. G., Kapan, D. D., Planas, S., Linares, M., Heckel, D. & McMillan, W. O. (2004) The historical demography of Müllerian mimicry in the neotropical *Heliconius* butterflies. *Proceedings of the National Academy of Sciences, USA*, **101**: 9704-9709.

Fox, R. M. (1940) A generic review the Ithomiinae (Lepidoptera: Nymphalidae). *Transactions of the American Entomologial Society*, **66** (1076): 161-207.

Freitas, A. V. L. & Brown, K. S. Jr. (2004) Phylogeny of the Nymphalidae (Lepidoptera). *Systematic Biology,* **53**: 363-383.

Gallusser, S., Guadagnuolo, R. & Rahier, M. (2004) Genetic (RAPD) diversity between *Oleria onega agarista* and *Oleria onega* ssp. (Ithomiinae, Nymphalidae, Lepidoptera) in north-eastern Peru. *Genetica,* **121**: 65-74.

Gascon, C., Malcolm, J. R., Patton, J. L., da Silva M. N. F., Bogart, J. P., Lougheed, S. C., Peres, C. A., Neckel, S. & Boag, P. T. (2000) Riverine barriers and the geographic distribution of Amazonian species. *Proceedings of the National Academy of Sciences, USA,* **97** (25): 13672-13677.

Gilbert, L. E. (2003) Adaptive novelty through introgression in *Heliconius* wing patterns: evidence for a shared genetic "Toolbox" from synthetic hybrid zones and a theory of diversification. In (Eds.) Boogs, C. L., Watt, W. B., Ehrlich, P. R. *Butterflies: ecology and evolution taking flight.* University of Chicago Press, Chicago and London.

Glor, R. E., Vitt, L. J. & Larson, A. (2001) A molecular phylogenetic analysis of diversification in Amazonian Anolis lizards. *Molecular Ecology,* **10** (11): 2661-2668.

Haber W.A. (1978) Evolutionary ecology of tropical mimetic butterflies (Lepidoptera: Ithomiinae). PhD thesis, University of Minnesota, USA.

Haffer, J. (1969) Speciation in amazonian forest birds. *Science,* **165**: 131-137.

Haffer, J. (1997) Alternative models of vertebrate speciation in Amazonia: a review. *Biodiversity and Conservation,* **6**: 451-476.

Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature (London),* **405**: 907-913.

Jiggins, C. D. & McMillan, W. O. (1997) The genetic basis of an adaptive radiation; mimicry in two *Heliconius* sibling species. *Proceedings of the Royal Society of London,* Series B, **264:** 1167-1175.

Jiggins, C. D., Naisbit R. E., Coe, R. L. & Mallet, J. (2001) Reproductive isolation caused by colour pattern mimicry. *Nature,* **411**: 302-305.

Jiggins, C. D., Mavarez, J., Beltrán, M., McMillan, W. O., Johnston J. S. & Bermingham, E. (2005) A genetic linkage map of the mimetic butterfly, *Heliconius melpomene. Genetics,* **171**: 557-570.

Jiggins, C. D., Mallarino, R., Willmott, K. W., Bermingham, E. (2006) Phylogenetic evidence for speciation caused by ecological adaptation in neotropical *Ithomia* butterflies (Lepidoptera; Nymphalidae). *Evolution,* **60** (7): 000-000.

Joron, M. & Mallet, J. (1998) Diversity in mimicry: paradox or paradigm? *Trends in Ecology and Evolution*, **13**: 461-466.

Joron, M., Wynne, I. R., Lamas, G. & Mallet, J. (1999) Variable selection and the coexistence of multiple mimetic forms of the butterfly *Heliconius numata*. *Evolutionary Ecology*, **13**: 721-754.

Lamas, G. (2004) Nymphalidae. Ithomiinae, pp. 172-191. In (Ed.) Lamas, G. Checklist: Part 4A. Hesperioidea - Papilionoidea. In (Ed.) Heppner, J. B. *Atlas of Neotropical Lepidoptera. Volume 5A.* Association for Tropical Lepidoptera, Scientific Publishers, Gainesville.

Lessa, E. P., Cook, J. A. & Patton, J. L. (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the late Quaternary. *Proceedings of the National Academy of Sciences, USA*, **100**: 10331-10334.

Lougheed, S. C., Gascon, C., Jones, D. A., Bogart, J. P. & Boag, P. T. (1999) Ridges and rivers: a test of competing hypotheses of Amazonian diversification using a dart-poison frog (*Epipedobates femoralis*). *Proceedings of the Royal Society of London, Series B*, **266** (1431): 1829-1835.

Lovette, I. J. (2004) Molecular phylogeny and plumage signal evolution in a trans Andean and circum Amazonian avian species complex. *Molecular Phylogenetics and Evolution*, **32** (2): 512-523.

Mallet, J. (1989) The genetics of warning colour in Peruvian hybrid zones of *Heliconius erato* and *H. melpomene*. *Proceedings of the Royal Society of London, Series B*, **236**: 163-85.

Mallet, J. (1993) Speciation, raciation, and color pattern evolution in *Heliconius* butterflies: evidence from hybrid zones. In (Ed.) Harrison, R. G. *Hybrid Zones and the Evolutionary Process*, pp. 226-260. New York: Oxford University Press.

Mallet, J. & Barton, N. H. (1989) Strong natural selection in a warning color hybrid zone. *Evolution*, **43**: 421-431.

Mallet, J. & Gilbert, L. E. (1995) Why are there so many mimicry rings? Correlations between habitat, behaviour and mimicry in *Heliconius* butterflies. *Biological Journal of the Linnean Society*, **55**: 159-180.

Mallet, J. & Joron, M. (1999) Evolution of diversity in warning color and mimicry: polymorphisms, shifting balance, and speciation. *Annual Review of Ecology and Systematics*, **30**: 201-233.

Mallet, J., McMillan, W. O. & Jiggins, C. D. (1998) Mimicry and warning color at the boundary between races and species. In (Ed.) Berlocher, S & Howard, D. *Endless Forms: Species and Speciation*, pp. 390-403. New York: Oxford University Press.

Masters, A. R. (1990) Pyrrolizidine alkaloids in artificial nectar protect adult ithomiine butterflies from a spider predator. *Biotropica*, **22** (3): 298-304.

Medina, M. C., Robbins, R. K. & Lamas, G. (1996) Vertical stratification of flight by ithomiine butterflies (Lepidoptera: Nymphalidae) at Pakitza, Manu National Park, Peru. In (Eds. Wilson, D. E & Sandoval, A. *Manu, The biodiversity of southeastern Peru*, pp 211-216. Washington, DC, Smithsonian Institution.

Müller, W. (1879) *Ituna* and *Thyridia*; a remarkable case of mimicry in butterflies. *Transactions of the Entomological Society of London*, **1879**: xx-xxix.

Naisbit, R. E., Jiggins, C. D. & Mallet, J. (2003) Mimicry: developmental genes that contribute to speciation. *Evolution and Development*, **5**: 269-280.

Papageorgis, C. (1975) Mimicry in neotropical butterflies. American Scientist, **63**: 522-532.

Patton, J. L. & Smith, M. F. (1992) MtDNA phylogeny of Andean mice: a test of diversification across ecological gradients. *Evolution*, **46** (1): 174-83.

Patton, J. L., da Silva, M. N. F., & Malcolm, J. R. (1994) Gene genealogy and differentiation among arboreal spiny rats (Rodentia, Echimyidae) of the Amazon basin: a test of the Riverine Barrier Hypothesis. *Evolution*, **48** (4): 1314-1323.

Patton, J. L., da Silva, M. N. F. & Malcolm. J. R. (2000) Mammals of the Rio Juruá and the evolutionary and ecological diversification of Amazonia. *Bulletin of the American Museum of Natural History*, **244**: 1-306.

Peres, C. A., Patton, J. L., & da Silva, M. N. F. (1996) Riverine barriers and gene flow in Amazonian saddle-back tamarins. *Folia Primatologica (Basel)* **67**:113–124.

Pliske, T. E. (1975) Attraction of Lepidoptera to plants containing pyrolizidine alkaloids. *Environmental Entomology*, **4**: 455-473.

Pough, F. H., Brower, L. P., Meck, H. R. & Kessel, S. R. (1973) Theoretical investigations of automimicry: multiple trial learning and the palatability spectrum. *Proceedings of the National Academy of Sciences of the United States of America*, **70**: 2261-2265.

Prum, R. O. (1988) Historical relationships among avian forest areas of the western Amazon. PhD dissertation, University of Florida. Ann Arbour University Microfilm International.

Schulz, S., Beccaloni, G., Brown, K. S. Jr, Boppre, M., Freitas, A. V. L., Ockenfels, P. & Trigo, J. R. (2004) Semiochemicals derived from pyrrolizidine alkaloids in male ithomiine butterflies. *Biochemical Systematics and Ecology*, **32** (8): 699-713.

Sheppard, P. M., Turner, J. R. G., Brown, K. S., Benson, W. W. & Singer, M. C. (1985) Genetics and the evolution of Müellerian mimicry in *Heliconius* butterflies. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **308**: 433-613.

Sick, H. (1967) Rios e enchentes na Amazônia como obstáculo para a avifauna. Atas. Simp. *Sôbre a Biota Amazônica*, **5** (Zoologia): 495-520.

Silva, K. L. & Trigo, J. R. (2002) Structure-activity relationships of pyrrolizidine alkaloids in insect chemical defense against the orb-weaving spider *Nephila clavipes*. *Journal of Chemical Ecology*, **28** (4): 657-668.

Svenning, J. C. (2001) On the role of microenvironmental heterogeneity in the ecology and diversification of neotropical rain-forest palms (Arecaceae). *Botanical Review*, **67** (1): 1-53.

Tobler, A., Kapan, D., Flanagan, N. S., Gonzalez, C., Peterson, E., Jiggins, C. D., Johntson, J. S., Heckel, D. G. & McMillan, W. O. (2005) First-generation linkage map of the warningly colored butterfly *Heliconius erato*. *Heredity*, **94** (4): 408-417.

Trigo, J. R., Barata, L. E. S. & Brown, K. S. (1994) Stereochemical inversion of pyrrolizidine alkaloids by *Mechanitis polymnia* (Lepidoptera, Nymphalidae, Ithomiinae)- Specificity and evolutionary significance. *Journal of Chemical Ecology*, **20** (11): 2883-2899.

Trigo, J. R., Brown, K. S. Jr., Witte, L., Hartmann, T., Ernst, L. & Barata, L. E. S. (1996) Pyrrolizidine alkaloids: different acquisition and use patterns in Apocynaceae and Solanaceae feeding ithomiine butterflies (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, **58**: 99-123.

Turner, J. R. G. (1971) Two thousand generations of hybridization in a *Heliconius* butterfly. *Evolution*, **25**: 471-482.

Turner, J. R. G. & Mallet, J. L. B. (1996) Did forest islands drive the diversity of warningly coloured butterflies? Biotic drift and the shifting balance. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **351**: 835-845.

Vasconcellos-Neto, J. & Lewinsohn, T. M. (1984) Discrimination and release of unpalatable butterflies by *Nephila clavipes*, a neotropical orb-weaving spider. *Ecological Entomology*, **9**: 337-344.

Wallace A. R. (1853) *A Narrative of travels on the Amazon and Rio Negra, with an account of the native tribes, and observations on the climate, geology and natural history of the Amazon valley.* Reeve & Co. London.

Wickler, W. (1968) *Mimicry in plants and animals.* McGraw-Hill New York.

Wiley, E. O. (1988) Vicariance biogeography. *Annual Review of Ecology and Systematics*, **19**: 513-542.

Willmott, K. R. & Mallet, J. (2004) Correlations between adult mimicry and larval host plants in ithomiine butterflies. *Proceedings of the Royal Society of London, Series B (Suppl.)*, **271**: S266-269.

Wuster, W., Ferguson, J. E., Quijada-Mascarenas, J. A., Pook, C. E., Salomao, M. D. & Thorpe, R. S. (2005) Tracing an invasion: landbridges, refugia, and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: *Crotalus durissus*). *Molecular Ecology*, **14** (4): 1095-1108.

CHAPTER TWO

STRIKINGLY VARIABLE DIVERGENCE TIMES INFERRED ACROSS AN AMAZONIAN BUTTERFLY 'SUTURE ZONE'

Abstract

'Suture zones' are areas where hybrid and contact zones of multiple taxa are clustered. Such zones have been regarded as strong evidence for allopatric divergence by proponents of the Pleistocene forest refugia theory, a vicariance hypothesis frequently used to explain diversification in the Amazon basin. A central prediction of the refugia and other vicariance theories is that the taxa should have a common history; that is divergence times should be coincident among taxa. A suture zone for Ithomiinae butterflies near Tarapoto, N. E. Peru, was therefore studied to examine divergence times of taxa in contact across the zone. We sequenced 1619 bp of the mitochondrial COI/COII region in 172 individuals of 31 species from across the suture zone. Inferred divergence times differed remarkably, with divergence between some pairs of widespread species (each of which may have two or more subspecies interacting in the zone, as in the genus *Melinaea*) being considerably less than that between hybridizing subspecies in other genera (for instance in *Oleria* spp.). Our data therefore strongly refute a simple hypothesis of simultaneous vicariance and suggest that ongoing parapatric or other modes of differentiation in continuous forest may be important in driving diversification in Amazonia.

Introduction

Various hypotheses have been proposed to explain the record richness of Amazonian biodiversity, of which the Pleistocene refuge hypothesis has received particular attention. The refuge hypothesis (Haffer 1969, Simpson & Haffer 1978, Haffer 1997, Haffer & Prance 2001) contends that dry intervals associated with climatic oscillations of the Tertiary and Quaternary reduced forest habitats to fragments isolated by expanses of open vegetation. Within such fragments evolutionary innovations accumulated, splitting phyletic lineages. By the time refuges re-expanded and brought forest biotas back into contact, the diverged taxa were hypothesised to have become partially or completely reproductively isolated.

Temperate glacial refuges are well documented and widely accepted (Hewitt 2000) and genetic studies based on Northern Hemisphere taxa have provided support for a temperate zone refuge hypothesis (Jaarola & Searle 2002). Pleistocene rainforest refuges have also been implicated in the historical isolation of tropical taxa (Schneider & Moritz 1999). However, a study which identified genetic evidence of the population growth associated with habitat expansions after the late Pleistocene in North American mammals, failed to do so for Amazonian small mammals (Lessa et al. 2003). In another recent study, Flanagan and colleagues (2004) reported that the radiation of wing patterning in the Neotropical *Heliconius erato* and *Heliconius melpomene* had not been coincident, and that these butterflies did not share historical demographies, findings which were also inconsistent with hypotheses of Pleistocene population fragmentation. Such reports are in accordance with recent paleoecological data from the Amazon basin (Colinvaux 1997, Colinvaux & De Oliveira 2000, Colinvaux et al. 2000, Colinvaux & De Oliveira 2001) which suggest that the climatic oscillations of the Pleistocene did not have a dramatic effect on forest cover. Using palynological analyses, Colinvaux and colleagues (Colinvaux & De Oliveira 2001, Colinvaux et al. 2001) identified species turnover in the flora, but not the biome replacements required to support historical forest fragmentation. These geological findings place Pleistocene refuge theory as an explanation for Neotropical biodiversity under strong attack.

Here, we test a key biological prediction of the Pleistocene refuge theory. Implicit in the theory, as well as in other vicariance theories, is that fragmentation would have had a simultaneous impact across the whole biota. Thus, lineage splits caused by forest fragmentation should have occurred at the same time in co-

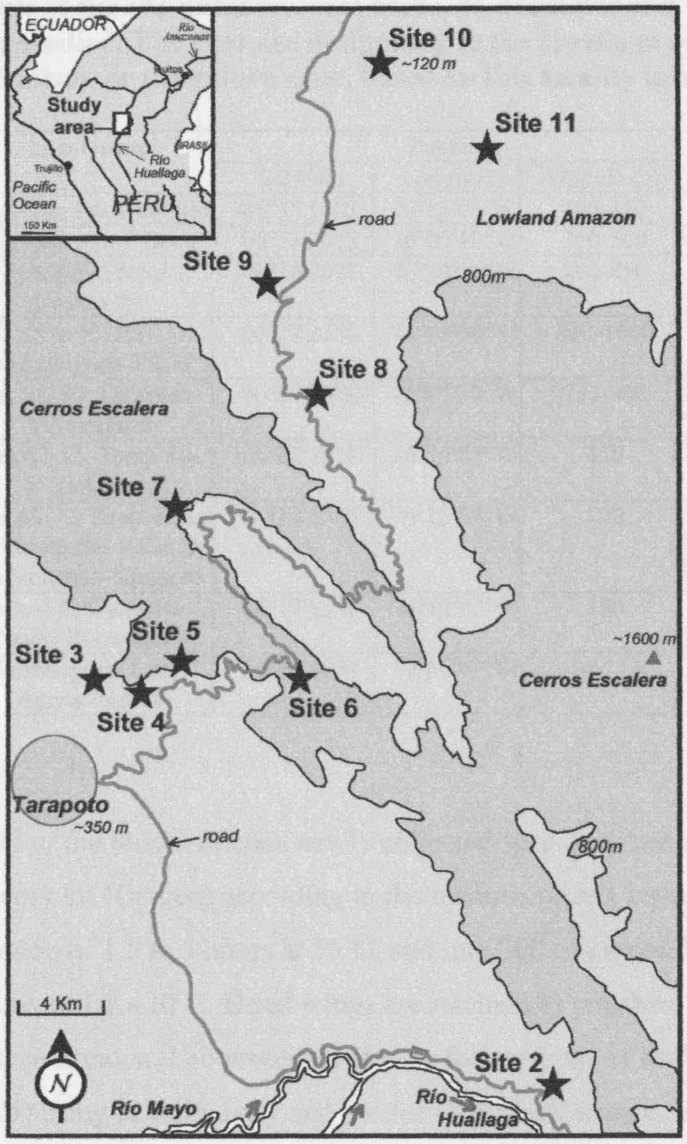
distributed taxa, and, in the case of the Pleistocene refuge theory, commensurate with Pleistocene climatic fluctuations. Clustered divergences due to vicariant events have previously been demonstrated, for example in marine species pairs found across the Isthmus of Panama (Knowlton et al. 1993, Marko 2002). In accordance with predictions of vicariance theories, lineage splits should be most strikingly correlated in pairs of closely related taxa, which share natural history traits and therefore can be expected to respond similarly to common environmentally-induced vicariance.

We here focus on the subfamily Ithomiinae (Lepidoptera: Nymphalidae). The group contains about 355 species (Lamas et al. 2004), each of which may consist of multiple strongly-differentiated subspecies, providing an opportunity to study multiple, co-distributed and closely related taxa. Each subspecies is typically involved in Müllerian mimicry 'rings' with sympatric taxa from many other species (Brown 1979, Joron & Mallet 1998). Additionally, the distribution of ithomiine subspecies has provided key biogeographic evidence for the refuge hypothesis (Haffer 1969, Brown 1979, Brown 1982, Brown 1987b, Turner & Mallet 1996).

We sampled from two adjacent centres of endemism in N.E. Peru separated by the Cerros Escalera mountain range (ranging up to ~1600m above sea level): the Río Mayo/upper Río Huallaga valley, and the lower Río Huallaga basin (Figure 1). These centres of endemism (Lamas 1982) were interpreted by Brown (1987a, 1987b) as evidence for the Huallaga (Andean valley) and Ucayali (Amazon basin) refuges, respectively. About 40 species of ithomiines show morphologically differentiated pairs of subspecies between these two endemic areas; their strikingly correlated contact and hybrid zones form a very well-defined 'suture zone' (*sensu* Remington 1968). Additionally, approx. 40 species are widespread and monomorphic, and found in both areas, while an additional approx. 40 monomorphic species are distributed only in the Huallaga centre, and approx. 20 are restricted to the Ucayali centre.

We investigate questions of ithomiine diversification using a 1619 bp aligned COI-tRNA-COII mitochondrial region as it has a high substitution rate, and is expected to show marked differences between even recently diverged taxa. We use a "rate-smoothing" method to investigate the relative chronology of diversifications.

Figure 1. Map of the study area in N. E. Peru (from Joron et al. 1999). Study site 12 is approximately 70 km N. W. from the top left hand corner of the map. See Table 1 for additional details about each locality studied.



Materials and methods

Ithomiine samples for DNA analysis were collected and preserved at UCL in 20% dimethylsulphoxide, 0.25M EDTA and saturated NaCl solution. Taxa were collected from study localities described by Joron and colleagues (1999), with the addition of site 12 (see Fig. 1 and Table 1). Samples were identified by Keith Willmott and Gerardo Lamas. Study taxa were selected at a range of taxonomic levels to obtain information spanning the whole process of differentiation. Where possible, at least three representatives of each species, from each centre of endemism, were sampled. For example; 3 *Hyposcada anchiala*

interrupta from site 11, and 3 *Hyposcada anchiala mendax* from site 12; and 3 *Ithomia agnosia agnosia* from both sites 2 and 12.

Table 1. Details of the site number, local name, co-ordinates and altitude of each locality studied. The sites are designated to the Huallaga or Ucayali zones of endemism, or the suture zone, based on this locality information.

Site number	Local name	Position			Zone of endemism
		Latitude	Longitude	Altitude (m)	
2	Quebrada Pucayaquillo	06°35'10"S	76°13'05"W	300-450	Huallaga (Andean)
3	Río Shilcayo + Boca	06°27'20"S	76°20'40"W	350-500	
4	Km 8-9. Sector Uruhuasha	06°28'00"S	76°20'05"W	700-850	
6	Km 15-19. Ahuashiyacu-Túnel	06°27'25"S	76°18'00"W	800-1000	
7	Km 28-30. Tarapoto-Yurimaguas	06°24'30"S	76°19'30"W	750-900	Suture Zone
8	Km 42-52. Tarapoto-Yurimaguas	06°22'10"S	76°16'45"W	450	
10	Km 68-72. Santa Rosa de Davidcillo + Km 26 Yurimaguas-Tarapoto	06°14'55"S	76°15'50"W	120	Ucayali (Amazonian)
11	Km 4-8. Carretera Pongo-Barranquita	06°17'15"S	76°13'40"W	150	
12	Serranoyacu + Pte. Aguas Verdes	05°67'49"S	77°67'97"W	980-1150	Huallaga (Andean)

One third of the thorax of each newly collected specimen was extracted using the DNAeasy kit (Qiagen) according to the manufacturer's instructions, with an initial incubation of 1.5 to 3 hours at 55°C, and into 300 µl elution buffer. Samples are preserved at –20°C. Dried wings are retained as vouchers at University College London. Lab protocols closely follow those of Mallarino and colleagues (2005) using primers Jerry and Pat for COI amplification and sequencing, George and Imelda for COII amplification, and George and Imelda, or Phyllis, Imelda and an additional primer, Romeo, (5'-TAATATGACAGATTATATGTAATGGA-3') for COII sequencing. Some *Pseudoscada* individuals would not amplify for COI, so a new specific primer (Talma, 5'- AATCAGAATAACGACGAGG -3') was developed, which successfully paired with Jerry, following the above protocol. Sequence data have been lodged in Genbank (DQ078312- 473; DQ078478-479). Sequence data for 8 specimens of *Ithomia* (AY713067-8, AY713040-1, and AY713075-8, Mallarino et al. 2005) and a danaine outgroup *Anetia briarea* (DQ071866, Whinnett et al. 2005) were also used. Vouchers for these specimens are retained by C. Jiggins at

The University of Edinburgh (*Ithomia*) and A. V. Z. Brower at Oregon State University (*Anetia*).

PAUP* version 4.0b 10 (Swofford 2000) was used to calculate the number of variable sites. Bayesian analysis was performed using MrBayes 3.0 (Huelsenbeck & Ronquist 2001), using a gamma rates heterogeneity model, in which all 6 reversible substitution types had different rates (nst=6). 4 simultaneous chains were run for 1,000,000 generations, and a tree was sampled every 100 generations. The tree with maximum posterior probability was assessed using a consensus of the final 9,000 trees (representing the final 900,000 generations), after confirmation that likelihood values stabilised after the first 100,000 generations. Pairwise divergences between taxa, corrected for multiple hits, were calculated in PAUP using maximum likelihood parameters identified using PAUP* with a variant of the Modeltest script (Posada & Crandall 1998); however, the Bayesian consensus tree was used, rather than a NJ tree as in the usual Modeltest script. Model parameters found and used subsequently were; base frequencies, A: 0.3537, T: 0.4764, C: 0.0788, G: 0.0911; relative transition rates, A↔C: 9.4363, A↔G: 26.4378, A↔T: 5.5769, C↔G: 4.7678 C↔T: 109.8022, and G↔T: 1.0000; gamma distributed rates with shape parameter: 0.6605; and fraction of invariable sites: 0.5626 (corresponding to a General Time Reversible, GTR+I+G model). We also used a PAUP* likelihood ratio method to test constancy of evolutionary rates (i.e. the molecular clock null hypothesis) across the Bayesian consensus tree.

Because the molecular clock was rejected, we used Yang's AHRS likelihood method in PAML (using the BASEML subprogram, see Yang 2004, Yang & Yoder 2003, Yoder & Yang 2000) to obtain branch lengths by fitting a discrete model of rate variation, implemented with four different rates chosen via a clustering analysis (Yang 2004), to the Bayesian consensus tree. There are no fossil data for Ithomiinae, so times of divergence were calculated relative to the Ithomiinae + *Anetia briarea* (Danainae) root, which was set to 1.0. Relative divergence times were then calibrated to molecular divergence by performing a linear regression of selected, well-supported deeper nodes of species within genera against GTR+I+G distances obtained in PAUP*. As the fit was good ($r^2 = 0.926$), we used node-depths instead of the more usual averaged pairwise divergences between taxa as estimators of the times of divergence between taxa. The resultant

tree therefore gives relative estimates of times of divergence in terms of fitted GTR+I+G distances.

Results

Sequence data were obtained for 172 individuals representing 31 species from 10 genera, plus an *Anetia* outgroup. The final 1619 bp alignment spanned 829 bp of the 3' end of COI (alignment positions 1-829), 64 bp tRNA (830-893) and 726 bp COII (894-1619) (see Fig. 2). Excluding *Anetia*, 532 sites were variable (32.9%), of which 511 (31.6%) were parsimony-informative. Likelihood ratio tests were used to test the molecular clock by investigating rate heterogeneity across the Bayesian consensus tree (shown in Fig. 3). Substitution rates were heterogeneous (2 ΔlnL = -205.95, 171 df, P = 0.031), implying a poor fit to the molecular clock.

The rate-smoothed tree obtained via MrBayes and PAML is shown in Fig. 4. Intergeneric relationships were not as predicted from morphology or molecular phylogenetic analyses using a panel of mitochondrial and nuclear genes (Brower et al. 2005). This is attributable to saturation of the mitochondrial data, and resultant weak support for deep branches in comparisons among phylogenetically distant taxa. However, all genera formed very strongly supported monophyletic clades (posterior probabilities of 1), with the exception of *Hypothyris*, which contained the genus *Hyalyris* due to the basal position of *Hypothyris cantobrica*, as well as due to the closer relationship of *Hypothyris* sp. nov. to *Hyalyris oulita* than to other *Hypothyris*. *H. cantobrica* was formerly placed in the monotypic genus *Rhodussa* and is notably divergent from remaining *Hypothyris* + *Hyalyris*, suggesting that *Rhodussa* might be a valid taxon if *Hyalyris* is to be kept separate. Apart from these, only *Oleria* was recovered as a genus with a weaker posterior probability, 0.87, likely due to the effects of saturation on this deep clade.

Within genera, haplotypes from the same species tended to cluster together. A striking exception is in the genus *Melinaea*. *M. marsaeus*, *M. menophilus*, *M. satevis* and *M. isocomma* haplotypes are clearly not clustered by species. In addition, subspecies of *Oleria onega* and *Brevioleria aelia* do not group together, instead clustering most closely with haplotypes from a different species. For the purposes of the current paper, we therefore consider the subspecies of *B. aelia* as separate species. However, the subspecies of *O. onega*, which produce morphologically identifiable hybrids in contact zones (such as

Figure 2. Mitochondrial sequence data for *Oleria onega janarilla* (02-1583) as in the final alignment, showing the 829 bp COI (upper case, black), 64 bp tRNA (lower case, boxed), and 726 bp COII (upper case, blue), gene regions.

------TGAAGTTTACATTTTAATTCTCCCCGGGTTTGGAATAATTTCTCATATTATTTCTCAAG
AAAGAGGAAAAAAAGAAACTTTTGGATCTTTAGGAATAATTTATGCTATAATGGCAATTGGACTT
TTAGGATTTATTGTTTGAGCTCATCATATATTCACAGTAGGAATAGATATTGATACCCGAGCTTA
TTTTACTTCAGCTACAATAATTATTGCTGTCCCAACTGGAATTAAAATTTTTAGTTGATTAGCTA
CCTTCCACGGAACACAAATTAATTATAGTCCTTCAATTTTATGAAGACTAGGATTTGTTTTTTTA
TTTACCGTAGGAGGATTAACAGGAGTAATCTTGGCAAATTCTTCAATTGACATTACTCTTCATGA
TACTTATTATGTTGTTGCTCATTTTCATTATGTTCTTTCTATAGGAGCAGTATTTGCTATTTTAG
GAGGATTTATTCATTGATATCCTTTATTTACAGGATTAACTTTAAATCCATATTTATTAAAAATT
CAATTTATTACTATATTTTTAGGAGTTAATATTACTTTTTTTCCTCAACATTTTTTAGGTTTAGC
CGGAATACCTCGTCGTTATTCAGATTATCCTGATAGATTCATTTCTTGAAATATTATCTCATCTT
TCGGGTCTTATATCTCGCTTTTATCAATAATAATAATTATTATTATTATTTGAGAATCTATAATT
AATCAACGAATTATTTTATTTTCATTAAATATACCATCATCTATTGAATGATATCAAAATTTACC
TCCTGCAGAACATTCTTATAATGAACTTCCCATCTTAAGTAATTTCTAAtatggcagattatatg
taatggatttaaaccccatttataaaggagttatccttttttagaa-ATGGCAACATGATCTAA
TTTTAATTTTCAAAATAGAGCTTCTCCTTTAATAGAACAAATTATTTTCTTTCATGATCATACTT
TAATTATTTTAATTATAATTACTATTTTAATTTCATATTTAATAATTAATTTATTTTTCAATAAA
TACACAAACCGATTTTTAATTGAAGAACAAATAATTGAATTAATTTGAACAATTTTACCAGCTTT
CACTTTAATTTTTATTGCTTTACCTTCGCTTCGACTATTATATCTTTTAGATGAAATTAATAATC
CTTTAATTACTTTAAAATCAATTGGTCATCAATGATATTGAAGTTATGAATACTCTGATTTTTTT
AATATCGAATTCGATTCTTATATAATTCAATCCTCTAATAATATTAAT------AATTTTCGATT
ATTAGATGTTGATAATCGTATTATTTTACCTATAAATAACCAAATTCGTATTTTAGTTACTGCTA
CAGATGTTATTCATTCATGAACAATCCCTGCTTTAGGAGTAAAAATTGATGGAAACCCAGGCCGA
TTAAATCAAACTAGGTTATTTATTAATCGGCCTGGATTATTTTTTGGTCAATGCTCAGAAATTTG
TGGAGCTAATCATAGCTTTATACCTATTGTAATTGAAAGTATCCCTATAAAAAACTTTATAGGAT
GAATTAATAATTATTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA

specimen 02-1958, see also Gallusser et al. 2004) are still considered to belong to *O. onega*, in spite of the evidence for paraphyly using these mitochondrial DNA data.

Many subspecies clusters are clearly recovered, as exemplified by the strong differentiation of both subspecies pairs in the genus *Hyposcada*. However, some well supported groupings of haplotypes conflict with their taxonomic assignment. It is possible that the position of *Pseudoscada florula gracilis* 1914 is the result of hybridisation and introgression, as wild hybrids are known between *P. florula gracilis* and *Pseudoscada florula aureola*. Additional cases of

individuals not clustering with other representatives from the same subspecies are; *Ithomia salapia aquinia* 72, *Napeogenes inachia ssp. nov.* 1188, *Brevioleria arzalia arzalia* 1203, and *Oleria gunilla lota* 209.

We found a remarkable range of pairwise comparisons between taxa, over the whole range of taxonomic levels (see Tables 2, 3, and Figs. 4, 5). Between-subspecies smoothed pairwise divergences ranged from 0.06% between *Melinaea menophilus* subspecies, and 6.40% between *Oleria onega* subspecies. Between-species divergences within genera varied by more than an order of magnitude, from 0.23% between *Melinaea menophilus* and *Melinaea isocomma*, to higher values such as 13.28% between *Hypothyris cantobrica* and the rest of *Hypothyris*, while divergences among genera within the Ithomiinae varied between 13.24% between *Pseudoscada* and *Brevioleria* to 17.74% at the base of the Ithomiinae.

Figure 3. Phylogenetic hypothesis based on mitochondrial nucleotide data for *Dircenna* (D), *Hyalyris* (Hl), *Brevioleria* (B), *Hyposcada* (Hs), *Hypothyris* (Ht), *Ithomia* (I), *Melinaea* (M), *Napeogenes* (N), *Oleria* (O) and *Pseudoscada* (P). *A. briarea* represents outgroup taxa. Tree topology and branch lengths are both the consensus of the last 9000 trees inferred using Bayesian methods. Branches with: Bayesian posterior probabilities (bpp) of 1 are red; bpp of 0.950-0.999 are blue; and bpp of 0.700-0.949 are yellow; all others are black. • indicates those specimens from the Huallaga zone, • the Ucayali zone, and • the suture zone. Site information for *Hyalyris oulita* 02-936 is missing.

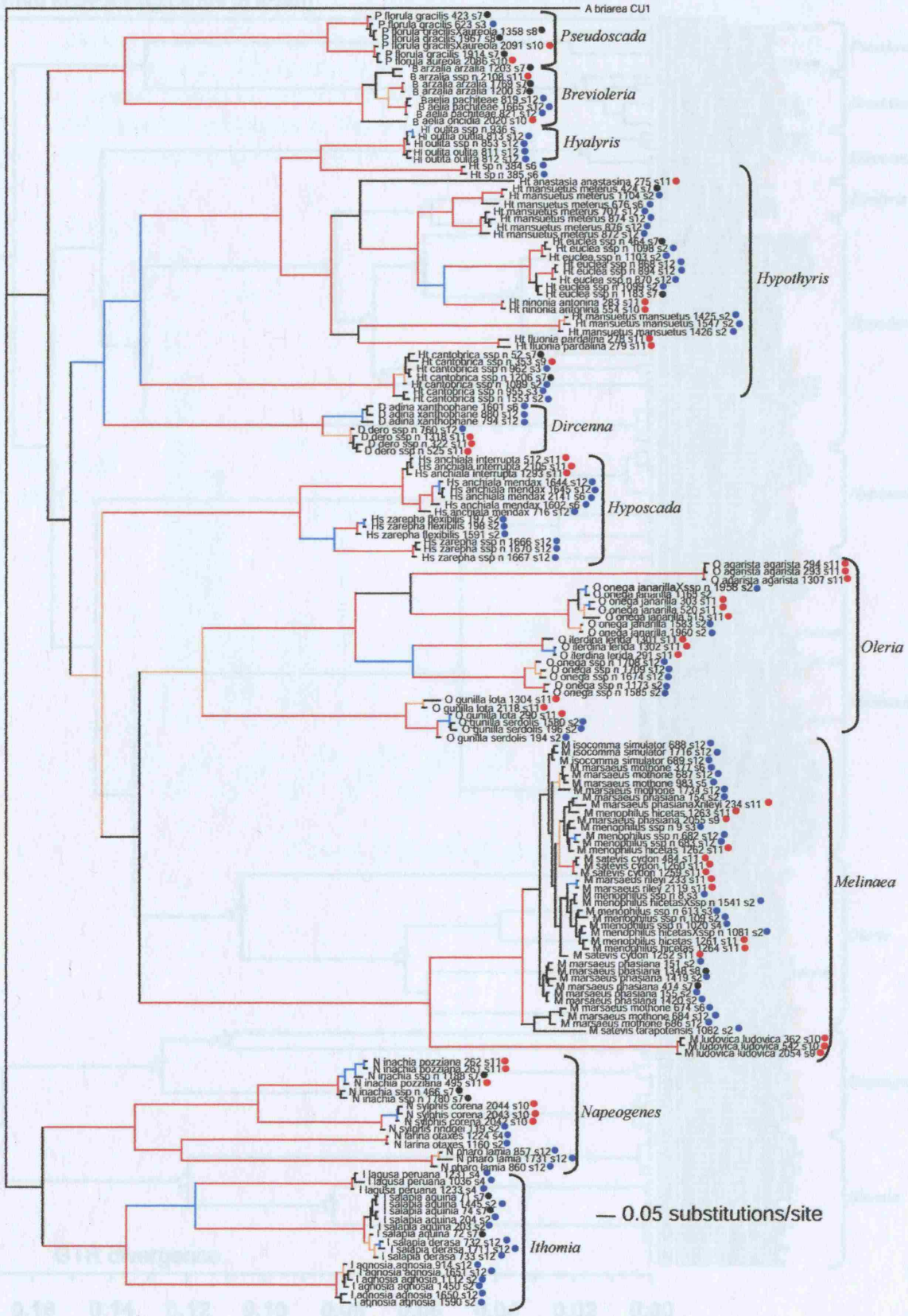


Figure 4. Phylogenetic hypothesis based on mitochondrial nucleotide data for *Dircenna* (D), *Hyalyris* (Hl), *Brevioleria* (B), *Hyposcada* (Hs), *Hypothyris* (Ht), *Ithomia* (I), *Melinaea* (M), *Napeogenes* (N), *Oleria* (O) and *Pseudoscada* (P). *A. briarea* represents outgroup taxa. Tree topology is consensus of the last 9000 trees inferred using Bayesian methods; branch lengths shown are obtained by the AHRS rate-smoothing procedure of Yang (2004), calibrated by a least squares fit to average GTR+I+G distances. • indicates those specimens from the Huallaga zone, • the Ucayali zone, and • the suture zone. Site information for *Hyalyris oulita* 02-936 is missing. ☆ indicates the nodes included in between-species comparisons in Fig. 5.

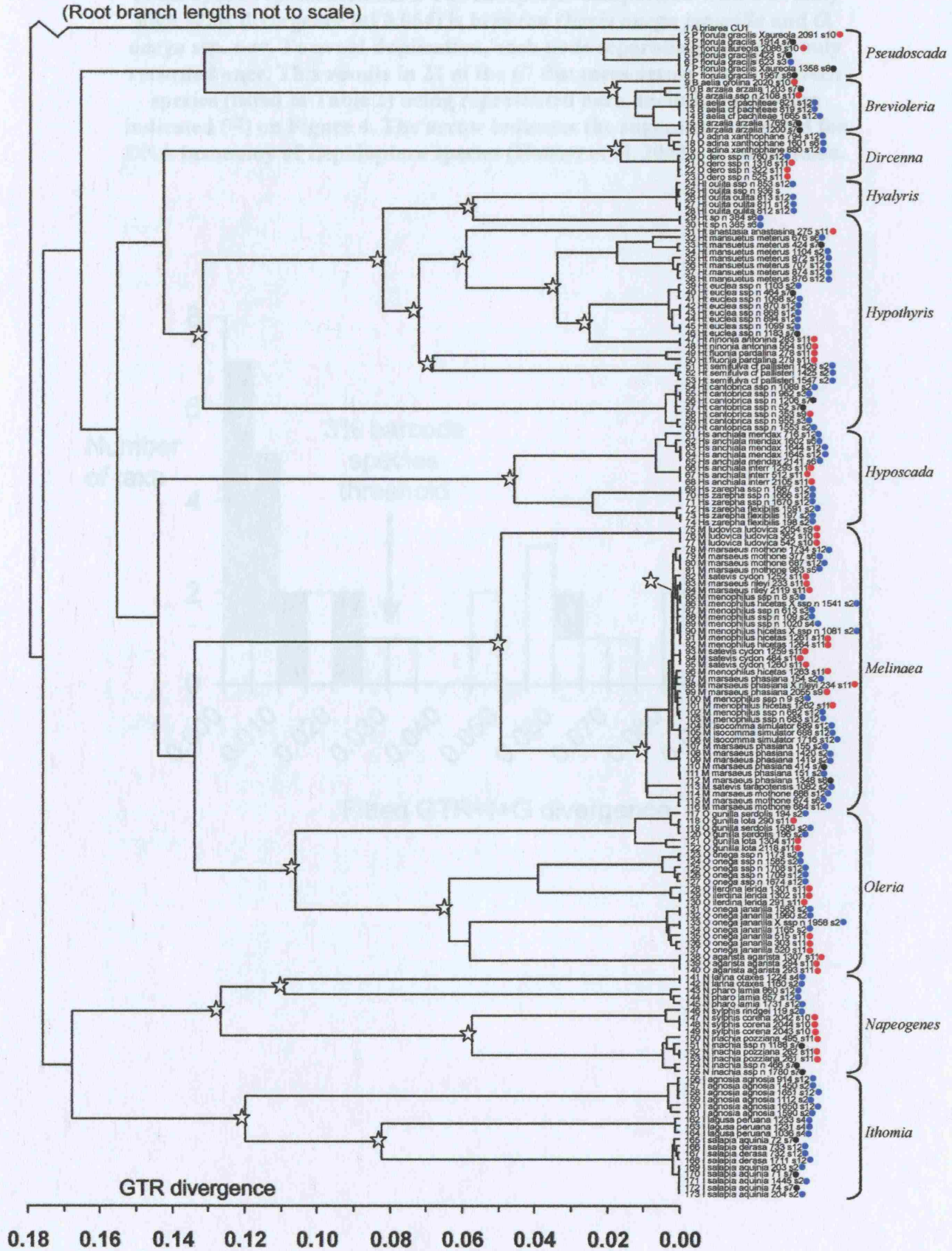


Figure 5. Histogram of fitted, within-genera GTR+I+G distances, across nodes between subspecies (black bars) and across nodes between species (white bars), in the rate-smoothed tree (Fig. 4). These distances are similar to (and calibrated using) traditionally measured average pairwise between taxon GTR+I+G distances. All 15 between subspecies comparisons (listed in Table 3) are represented here. The subspecies comparison with a notably high GTR divergence (of 0.064) is between *Oleria onega janarilla* and *O. onega* ssp. nov. To avoid duplication, each node separating species is only recorded once. This results in 21 of the 67 distances across nodes between species (listed in Table 2) being represented here. Included nodes are indicated (☆) on Figure 4. The arrow indicates the suggested threshold for DNA taxonomy of Lepidoptera species (Hebert et al. 2003), see discussion.

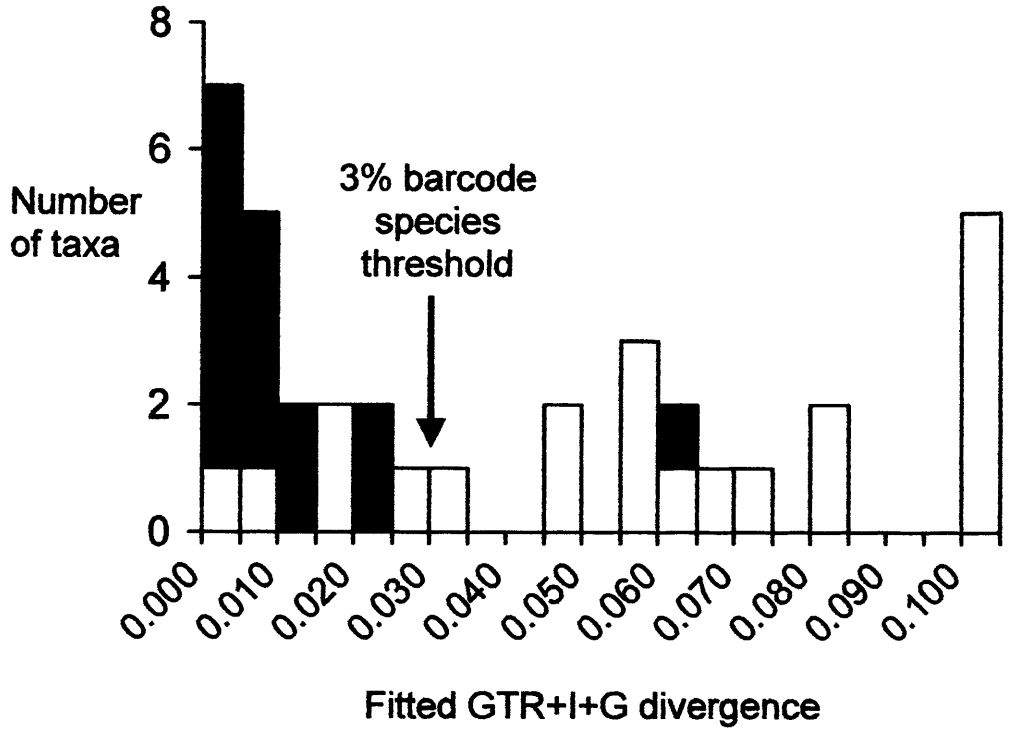


Table 2. Between-species pairwise % divergences. Divergences were calculated by calibrating the nodes of the rate smoothed tree (Fig. 4) to the GTR+I+G distances obtained in PAUP*, using a least squares fit. These values represent the distance to the deepest node separating, and therefore greatest divergence between, the two species. Species within a genus are allocated a species number (#), for the purpose of identifying which two species are being compared. For example, *Hypothyris euclea* is allocated the species number 3, and *Hypothyris fluonia* the species number 4 (as indicated by the digit in the # column, immediately to the right of the species name). The cell in the *Hypothyris* block corresponding to the compared species numbers 3 and 4 (i.e. 4 in the # column, and 3 in the # row, containing the value 7.24 in red text) provides the % divergence between *H. euclea* and *H. fluonia*.

Genus	Species	Between species pairwise comparisons									
		#	1	2	3	4	5	6	7	8	9
Dircenna	*adina*	1	X								
	dero	2	**1.65**	X							
Brevioleria	*aelia orolina*	1	X								
	aelia cf. *pachiteae*	2	**1.75**	X							
	arzalia	3	**1.75**	**1.23**	X						
Hyposcada	*anchiala*	1	X								
	zarepha	2	**4.61**	X							
Hypothyris	*anastasia*	1	X								
	cantobrica	2	**13.28**	X							
	euclea	3	**7.24**	**13.28**	X						
	fluonia	4	**7.24**	**13.28**	**7.24**	X					
	semifulva cf. *pallisteri*	5	**7.24**	**13.28**	**7.24**	**6.86**	X				
	mansuetus. meterus	6	**5.93**	**13.28**	**3.40**	**7.24**	**7.24**	X			
	ninonia	7	**7.24**	**13.28**	**2.52**	**7.24**	**7.24**	**3.40**	X		
	sp. n.	8	**8.25**	**13.28**	**8.25**	**8.25**	**8.25**	**8.25**	**8.25**	X	
Hyalyris	*oulita*	9	**8.25**	**13.28**	**8.25**	**8.25**	**8.25**	**8.25**	**8.25**	**5.74**	X
Ithomia	*agnosia*	1	X								
	lagusa	2	**12.08**	X							
	salapia	3	**12.08**	**8.29**	X						
Melinaea	*isocomma*	1	X								
	ludovica	2	**4.97**	X							
	marsaeus	3	**0.94**	**4.97**	X						
	menophilus	4	**0.23**	**4.97**	**0.94**	X					
	satevis	5	**0.94**	**4.97**	**0.94**	**0.94**	X				
Napeogenes	*inachia*	1	X								
	larina	2	**12.77**	X							
	pharo	3	**12.77**	**10.99**	X						
	sylphis	4	**5.76**	**12.77**	**12.77**	X					
Oleria	*agarista*	1	X								
	gunilla	2	**10.68**	X							
	ilerdina	3	**6.40**	**10.68**	X						
	onega	4	**6.40**	**10.68**	**6.40**	X					

Table 3. Between-subspecies pairwise % divergences. Divergences were calculated by calibrating the nodes of the rate smoothed tree (Fig. 4) to the GTR+I+G distances obtained in PAUP*, using a least squares fit. The values represent the distance to the deepest node separating, and therefore greatest divergence between, the two subspecies.

Subspecies	Pairwise % divergence
B. a. arzalia + ssp. n.	1.23
H. a. interrupta + *mendax*	0.68
H. z. flexibilis + ssp. n.	2.43
I. s. aquinia + *derasa*	0.23
M. m. mothone + *phasiana*	0.94
M. m. mothone + *rileyi*	0.94
M. m. phasiana + *rileyi*	0.31
M. m. hicetas + ssp. n.	0.06
M. s. cydon + *tarapotensis*	0.94
N. i. pozziana + ssp. n.	0.46
N. s. corena + *rindgei*	0.39
O. g. lota + *serdolis*	1.25
O. o. janarilla + ssp. n.	6.40
P. f. gracilis + *aureola*	2.11
H. o. oulita + ssp. n.	0.09

Discussion

Perhaps the major findings of this work are the highly variable levels of divergence in mtDNA between taxa of the same rank. Molecular divergence is an estimator of relative time of divergence, so these results seem to imply that times of divergence are also highly scattered. Our data did not conform to clock-like molecular evolution, so we estimated instead rate-smoothed percent divergences as a metric of time. (Forcing our tree to obey a molecular clock, however, made little difference to the overall finding of variability in times of divergence – data not shown). Applying Brower's (1994) 2.3% sequence divergence at COI and COII per million years to the rate-smoothed divergence values, subspecies within species, species with genera, and species among genera (including *Hyalyris* and *Hypothyris* in the same genus) appear to have diverged 0.03 – 2.78, 0.10 – 5.77, and 5.76 – 7.71 My ago, respectively. Zakharov and colleagues (2004) recently suggested substitution rates for COI and COII to be 0.78-1.02% sequence divergence per million years, much less than those calculated by Brower. Approximating this to 0.90%, this would increase our respective estimated times of divergence to 0.07 – 7.11, 0.26 – 14.76, and 14.71 – 19.71 My ago.

This variability of taxon divergences challenges assumptions underlying DNA "barcoding", which aims to categorise taxa as species on the basis of mtDNA divergence. Proponents of DNA barcoding have suggested a 3% threshold at the COI mitochondrial locus for Lepidoptera species (Hebert et al. 2003). However, six of our 23 species divergence nodes (within genera) had smoothed GTR distance values less than 3% (Fig. 5, Table 2), and one divergence node out of 16 between conspecific subspecies gave a smoothed distance value of greater than 3%. Somewhat later, Hebert and colleagues (2004) identified 10 cryptic species from within a single taxonomically recognized species of *Astraptes* (Hesperiidae), 8 of which had divergences < 3%, some of which were only 0.2-0.3% divergent (comparable to our closest *Melinaea*). *Heliconius* may have virtually no mtDNA divergence (0-2%) between subspecies, while values of >~2% were reported for closest species (Brower 1994, Brower & Egan 1997, Flanagan et al. 2004). Although the use of DNA as a tool in taxonomy is well-established, DNA taxonomy alone would have failed to recognise many of these species and/or would have recognized too many if the threshold was lowered to say 1% or 2%. Our results strongly argue against use of a 3% COI mtDNA 'barcode' (or any other threshold divergence value) as a protocol to distinguish species without supplementation by other methods.

The concordant distributions of subspecies differing in mimetic colour pattern demonstrate that the Cerros Escalera is a suture zone for many ithomiine and heliconiine taxa. By recovering repeated genealogical splits between multiple pairs of taxa distributed across this divide, our work shows the same pattern for molecular data. Coyne and Orr (2004) recently identified three requirements by which molecular data could provide evidence for vicariance (for example a Pleistocene refuge hypothesis). (1) Vicariance events should affect co-distributed taxa simultaneously, so that a main prediction of vicariance hypotheses is that divergence times should be similar and should correlate with the timing of relevant geological events, in this case in the Pleistocene. (2) Each interacting pair should consist of sister taxa. (3) There should be evidence of reproductive isolation between members of each pair.

Our results were in striking contrast to the first of these requirements. We found remarkable variability in pairwise divergences between different groups of taxa in comparisons at each taxonomic level. A number of hypotheses might

explain the observed variance in sequence divergence values. The first is variability of mutation rates and the controversial reliability of molecular clocks (Graur & Martin 2004). However, we believe that the method of rate-smoothing used here (Yang 2004) accommodates such rate variation to calibrate the divergence to time as well as or better than any other existing method.

It is also possible that the mitochondrial genealogies do not accurately reflect organismal tree topologies, due to confounding factors such as selection, nuclear copies or cytoplasmic factors (Hurst & Jiggins 2005), or because coalescence time variance is itself large (e.g. Hudson & Turelli 2003). To investigate this further, we have sequenced the nuclear loci, *triose phosphate isomerase*, *wingless* and *elongation factor 1-α*. Preliminary comparisons between two of the most strongly contrasting genera, *Oleria* and *Melinaea*, show molecular divergence in these nuclear regions that correlate with the strongly divergent patterns obtained using mtDNA data.

Furthermore, although some diversification seems to have occurred during the Pleistocene, molecular clock dating also shows genetic breaks both before and after this period. Therefore, ithomiine taxa do not at all seem to share the same clear genetic signature of subspecific or specific diversification due to one or a few historical events during the Pleistocene.

In their second requirement, Coyne and Orr (2004) argue that members of each pair should be sister taxa. It is very likely that at least some of the pairs of subspecies studied here are not sister taxa, but determining this is beyond the scope of this study, since it would involve the collection and sequencing of as many as 30 closely related subspecies per species from across the Neotropics. However, whether or not the pairs of subspecies are sister taxa should not affect the central prediction of vicariance for the suture zone under study here. All pairs of taxa across the suture zone should have the same (non-sister) relationships if generated by the same sequence of vicariance events, so that their branching pattern should be similar. Their divergence times should likewise be comparable. This prediction is clearly rejected. An alternative hypothesis therefore seems much more likely; the clustering of multiple contact zones in the suture zone must have been caused by range changes and/or independent evolution, rather than primarily by the same vicariance events in each group.

Finally, Coyne and Orr (2004) highlight the importance for evidence of reproductive isolation between members of each taxon pair. Past and present gene exchange, and persistent ancestral polymorphisms are hard to distinguish; these effects can affect deductions about the history of taxa. Here, the geographically differentiated ithomiines across the suture zone are classified as subspecies and hybridize to a greater or lesser extent. Putative hybrids, based on intermediacy of morphological characters, between subspecies are readily found, for example, *Oleria onega* ssp. nov. x *Oleria onega janarilla* (02-1958), and *Melinaea marsaeus phasiana* x *Melinaea marsaeus rileyi* (02-234). However, despite commonly hybridizing, the *Oleria onega* subspecies remain markedly genetically distinct. *Melinaea* are in stark contrast, with no genetic differences evident between subspecies. The species *M. isocomma, M. marsaeus, M. satevis* and *M. menophilus* also share identical or very similar haplotypes (mean between species divergences were all 0.9% or less), even though we found no evidence of hybridization between species. Nuclear data show a similar pattern (Zimmermann et al, pers. comm.). This is quite unexpected as experts have long recognised the distinctness of these *Melinaea* species (Brown 1979, Lamas et al. 2004), although there have been some revisions, name changes and reallocation of subspecies among species.

Morphologically distinct taxa showing high levels of gene exchange at many loci are predicted to be rare, so whilst we can expect some ancestral polymorphism and introgression between very closely related species, we hardly expect such effects among a group of four highly distinct and relatively widespread species (Brown 1977, Lamas et al. 2004), as found here. The four closely related *Melinaea* species comprise, between them, 34 morphologically recognizable subspecies, each of which typically has a divergent colour pattern affecting Müllerian mimicry (Brown 1977, Lamas et al. 2004). Our genetic data provides evidence for very recent diversification for the four closely related *Melinaea* species and even more recent diversification of their subspecies, coupled with incomplete lineage sorting. All of this diversification of *Melinaea* has apparently occurred well within the time that it took co-distributed pairs of subspecies in *Oleria onega*, O. *gunilla, Brevioleria arzalia, Hyposcada zarepha,* or *Pseudoscada florula* to diverge.

In conclusion, we infer strikingly varied divergence dates between pairs of taxa interacting in a single suture zone. This variability must be due to processes acting independently in different species and at different times, rather than due to a single or a few vicariant events. Our results imply that diversification of subspecies pairs across a single region are due neither to concordant evolutionary histories nor to simultaneous divergence times. We postulate an alternative: that ithomiine biogeographic patterns are explained by ongoing diversification through repeated and idiosyncratic parapatric or allopatric speciation in nearly continuous forest, interspersed by rapid range changes (Benson 1982, Mallet 1993, Knapp & Mallet 2003). The well-defined Escalera suture zone we see today may then be explained by redistribution of subspecific differentiation across ecotones and by partial environmental barriers to the movement of colour pattern and molecular genes. These current or recent factors must affect multiple pairs of taxa, in order for hybrid zones to collect together in the same suture zone (see also Mallet 1993). Our comparative data are therefore among the first to demonstrate that Amazonian diversification is relatively continuous, as would be likely if speciation took place via parapatric evolution in continuous forest.

References

Benson, W. W. (1982) Alternative models for infrageneric diversification in the humid tropics: tests with passion vine butterflies. In *Biological Diversification in the Tropics* (ed. G. T. Prance), pp. 608-640. New York, NY: Columbia Univ. Press.

Brower, A. V. Z. (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**: 6491-6495.

Brower, A. V. Z. & Egan, M. G. (1997) Cladistic analysis of *Heliconius* butterflies and relatives (Nymphalidae: Heliconiiti): a revised phylogenetic position for *Eueides* based on sequences from mitochondrial DNA and a nuclear gene. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **264**: 969-977.

Brower, A. V. Z., Freitas, A. V. L., Lee, M.-M., Silva Brandão, K. L., Whinnett A., & Willmott, K. R. (2006) Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions. *Systematic Entomology*, **31**: 288-301.

Brown, K. S. (1977) Geographical patterns of evolution in neotropical Lepidoptera: differentiation of the species of *Melinaea* and *Mechanitis* (Nymphalidae: Ithomiinae). *Systematic Entomology*, **2**: 161-197.

Brown, K. S. (1979) *Ecologia Geográfica e Evolução nas Florestas Neotropicais*. Campinas, Brazil: Universidade Estadual de Campinas.

Brown, K. S. (1982) Historical and ecological factors in the biogeography of aposematic Neotropical butterflies. *American Zoologist*, **22**: 453-471.

Brown, K. S. (1987a) Areas where humid tropical forest probably persisted. In *Biogeography and Quaternary History in Tropical America* (ed. T. C. Whitmore & G. T. Prance), pp. 45. Oxford, U.K.: Oxford Univ. Press.

Brown, K. S. (1987b) Biogeography and evolution of neotropical butterflies. In *Biogeography and Quaternary History in Tropical America* (ed. T. C. Whitmore & G. T. Prance), pp. 66-104. Oxford, U.K.: Oxford Univ. Press.

Colinvaux, P. A. (1997) An arid Amazon? *Trends in Ecology and Evolution*, **12**: 318-319.

Colinvaux, P. A. & De Oliveira, P. E. (2000) Palaeoecology and climate of the Amazon basin during the last glacial cycle. *Journal of Quaternary Science*, **15**: 347-356.

Colinvaux, P. A. & De Oliveira, P. E. (2001) Amazon plant diversity and climate through the Cenozoic. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **166**: 51-63.

Colinvaux, P. A., De Oliveira, P. E. & Bush, M. B. (2000) Amazonian and neotropical plant communities on glacial time-scales: The failure of the aridity and refuge hypotheses. *Quaternary Science Reviews*, **19**: 141-169.

Colinvaux, P. A., Irion, G., Rasanen, M. E., Bush, M. B. & De Mello, J. A. S. N. (2001) A paradigm to be discarded: Geological and paleoecological data falsify the Haffer & Prance refuge hypothesis of Amazonian speciation. *Amazoniana*, **16**: 606-607.

Coyne, J. A. & Orr, H. A. (2004) *Speciation*. Sunderland, Massachusetts: Sinauer Associates.

Flanagan, N. S., Tobler, A., Davison, A., Pybus, O. G., Kapan, D. D., Planas, S., Linares, M., Heckel, D. & McMillan, W. O. (2004) The historical demography of Müllerian mimicry in the neotropical *Heliconius* butterflies. *Proceedings of the National Academy of Sciences, USA*, **101**: 9704-9709.

Gallusser, S., Guadagnuolo, R. & Rahier, M. (2004) Genetic (RAPD) diversity between *Oleria onega agarista* and *Oleria onega* ssp. (Ithomiinae, Nymphalidae, Lepidoptera) in north-eastern Peru. *Genetica*, **121**: 65-74.

Graur, D. & Martin, W. (2004) Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics*, **20**: 80-86.

Haffer, J. (1969) Speciation in amazonian forest birds. *Science*, **165**: 131-137.

Haffer, J. (1997) Alternative models of vertebrate speciation in Amazonia: a review. *Biodiversity and Conservation*, **6**: 451-476.

Haffer, J. & Prance, G. T. (2001) Climatic forcing of evolution in Amazonia during the Cenozoic: on the refuge theory of biotic differentiation. *Amazoniana*, **16**: 579-607.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **270**: 313-321.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences, USA*, **101**: 14812-14817.

Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature (London)*, **405**: 907-913.

Hudson, R. R. & Turelli, M. (2003) Stochasticity overrules the "three-times rule": genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, **57**: 182-190.

Huelsenbeck, J. P. & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**: 754-755.

Hurst, G. D. D. & Jiggins, F. M. (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic, and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **272**: 1524-1534.

Jaarola, M. & Searle, J. B. (2002) Phylogeography of field voles (*Microtus agrestis*) in Eurasia inferred from mitochondrial DNA sequences. *Molecular Ecology*, **11**: 2613-2621.

Joron, M. & Mallet, J. (1998) Diversity in mimicry: paradox or paradigm? *Trends in Ecology and Evolution*, **13**: 461-466.

Joron, M., Wynne, I. R., Lamas, G. & Mallet, J. (1999) Variable selection and the coexistence of multiple mimetic forms of the butterfly *Heliconius numata*. *Evolutionary Ecology*, **13**: 721-754.

Knapp, S. & Mallet, J. (2003) Refuting refugia. *Science*, **300**: 71-72.

Knowlton, N., Weigt, L. A., Solorzano, L. A., Mills, D. K. & Bermingham, E. (1993) Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. *Science*, **260**: 1629-1632.

Lamas, G. (1982) A preliminary zoogeographical division of Peru based on butterfly distributions (Lepidoptera, Papilionoidea). In *Biological Diversification in the Tropics* (ed. G. T. Prance), pp. 336-357. New York: Columbia Univ. Press.

Lamas, G., Callaghan, C., Casagrande, M. M., Mielke, O., Pyrcz, T., Robbins, R. & Viloria, A. (2004) Hesperioidea and Papilionoidea. (Series Ed: Heppner, J.B. *Atlas of Neotropical Lepidoptera, Checklist*, Part **3**). Association for Tropical Lepidoptera/Scientific Publishers, Gainesville, Florida.

Lessa, E. P., Cook, J. A. & Patton, J. L. (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the late Quaternary. *Proceedings of the National Academy of Sciences, USA*, **100**: 10331-10334.

Mallarino, R., Bermingham, E., Willmott, K. R., Whinnett, A. & Jiggins, C. D. (2005) Molecular systematics of the butterfly genus *Ithomia* (Lepidoptera: Ithomiinae): a composite phylogenetic hypothesis based on seven genes. *Molecular Phylogenetics and Evolution*, **34**: 625-644.

Mallet, J. (1993) Speciation, raciation, and color pattern evolution in *Heliconius* butterflies: evidence from hybrid zones. In *Hybrid Zones and the Evolutionary Process* (ed. R. G. Harrison), pp. 226-260. New York: Oxford University Press.

Marko, P. B. (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology and Evolution*, **19**: 2005-2021.

Posada, N. M. & Crandall, K. A. (1998) *MODELTEST*: testing the model of DNA substitution. *Bioinformatics*, **14**: 817-818.

Remington, C. L. (1968) Suture-zones of hybrid interaction between recently joined biotas. *Evolutionary Biology*, **1**: 321-428.

Schneider, C. & Moritz, C. (1999) Rainforest refugia and evolution in Australia's Wet Tropics. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **266**: 191-196.

Simpson, B. B. & Haffer, J. (1978) Speciation patterns in the Amazonian forest biota. *Annual Review of Ecology and Systematics*, **9**: 497-518.

Swofford, D. L. (2000) *PAUP*: Phylogenetic Analysis Using Parsiomony (*and Other Methods)*. Sunderland, Massachusetts: Sinauer Associates.

Turner, J. R. G. & Mallet, J. L. B. (1996) Did forest islands drive the diversity of warningly coloured butterflies? Biotic drift and the shifting balance. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **351**: 835-845.

Whinnett, A., Brower, A. V. Z., Lee, M.-M., Willmott, K. R., Mallet, J. (2005) Phylogenetic utility of *Tektin*, a novel region for inferring systematic relationships among Lepidoptera. *Annals of the Entomological Society of America*, **98**: 873-886.

Yang, Z. (2004) A heuristic rate smoothing procedure for maximum likelihood estimation of species divergence times. *Acta Zoologica Sinica*, **50**: 645-656.

Yang, Z. & Yoder, A. D. (2003) Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemurspecies. *Systematic Biology*, **52**: 705-716.

Yoder, A. D. & Yang, Z. (2000) Estimation of primate speciation dates using local molecular clocks. *Molecular Biology and Evolution*, **17**: 1081-1090.

Zakharov, E. V., Caterino, M. S. & Sperling, F. A. H. (2004) Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). *Systematic Biology*, **53**: 193-215.

CHAPTER THREE

DIVERGENCES AT *TRIOSEPHOSPHATE ISOMERASE* (*Tpi*) AND MITOCHONDRIAL LOCI IN AMAZONIAN BUTTERFLIES: FURTHER INSIGHTS INTO NEOTROPICAL DIVERSIFICATION AND COMMENTS ON THE MOLECULAR EVOLUTION OF *Tpi*

Abstract

A central prediction of the Pleistocene forest refugia theory, a vicariance hypothesis frequently used to explain diversification in the Amazon basin, is that inferred divergence times should be coincidental among sympatric taxa. Using mitochondrial DNA (mtDNA) from 172 individuals of the butterfly subfamily Ithomiinae (Lepidoptera: Nymphalidae), we previously reported strikingly varied divergences at all taxonomic levels. This finding strongly rejected the simple hypothesis that simultaneous vicariance drove the diversification of all our studied taxa. However, the reliability of inferences based on mtDNA and single loci are frequently questioned and it was therefore of considerable interest to investigate if previous findings would be corroborated by additional data from a nuclear locus.

Here, we sequenced a 989-1144 bp region of *Triose-phosphate isomerase* (*Tpi*), corresponding to exons 1-5 of *Heliothis Tpi*, in 95 individuals of the Ithomiinae. Bayesian analysis recovered a branching order similar to that previously recovered with mtDNA. In addition, a marked variability in pairwise divergences for comparisons at different taxonomic levels was recovered with *Tpi*, mirroring the finding based on mtDNA. Although parallel divergences were recovered for some of the study taxa, evolution at mtDNA and *Tpi* was not always tightly correlated.

In addition, we discuss aspects of the molecular evolution of *Tpi*. We found that the *Tpi* intron evolves faster, and the *Tpi* exon slower, than mtDNA. Our data also showed that *Tpi* intron 1 evolved at approximately half the rate of the total *Tpi* intron data.

Introduction

Despite the wide application of mitochondrial DNA (mtDNA) in phylogenetic analysis, the reliability of inferences based on this source of data are frequently questioned. Some authors have reported non-neutral mtDNA evolution (Hudson & Turelli 2003, Ballard & Whitlock 2004), whilst others have reported the confounding effects of cytoplasmic factors, such as the intracellular bacterium *Wolbachia* (Hurst & Jiggins 2005), or that the maternal mode of inheritance might not accurately represent historical processes. In addition, potentially confounding nuclear copies that evolve at a different rate to their mitochondrial counterparts have been identified (Bensasson et al. 2001).

Discordances between loci are well documented. For example, in butterflies, Mallarino et al (2005) and Brower et al (2006) reported conflicting signals from mtDNA and nuclear loci in ithomiines. In another study, Beltrán et al (2002) and Bull (2003) presented evidence for gene flow in the autosomal loci *Mpi* and *Ci*, but not in mtDNA, sex-linked *Tpi* or autosomal *DDC* in *Heliconius*. Both theory, and the frequently reported cases of incongruent topologies, increasingly query the reliability of inferences derived from a single locus (Pamilo & Nei 1988, Edwards & Beerli 2000).

We previously investigated divergences between 172 ithomiine (Lepidoptera: Nymphalidae) taxa, sampled from two adjacent centres of endemism in N. E. Peru, using a 1619 bp region of mtDNA (Whinnett et al. 2005b). We found the taxa had strikingly varied divergences, for example, mean between-subspecies divergences calculated using maximum likelihood ranged from 0.18 % between *Melinaea menophilus* subspecies to 10.19 % between *Oleria onega* subspecies. This finding strongly rejected a simple hypothesis of a simultaneous vicariance driving the diversification of all taxa. It followed that the studied butterflies did not conform to one of the leading hypotheses used to explain the richness of Amazonian biodiversity, the Pleistocene Refuge theory. Given the above mtDNA and single marker concerns, we were interested in investigating whether our previous findings were an artefact of the chosen locus, or if the mtDNA results would be corroborated with data from a nuclear locus.

Triose-phosphate isomerase (*Tpi*) is an important enzyme for carbohydrate metabolism (Logsden et al. 1995). Using primers Tpi-1 and Tpi-2,

which are situated in exons 3 and 4 of *Heliothis Tpi* (U23080), Beltrán and colleagues (2002) amplified 155 bp of *Tpi* exon, plus 0-457 bp of *Tpi* intron in 33 *Heliconius* butterflies representing 21 species (AF413752-AF413797). *Tpi* was selected here based on their finding that it evolved rapidly enough for studies between geographic races of the same species. Few informative characters were identified at the within-species level in the region amplified with primers Tpi-1 and Tpi-2 in ithomiines, so we developed primers which amplify a longer region, corresponding to exons 1 and 5 in *Heliothis Tpi*.

In birds and butterflies, females are the heterogametic sex (ZW) and males the homogametic sex (ZZ). *Tpi* is located on the Z chromosome in *Heliconius* (Jiggins et al. 2005), and if similarly located in ithomiines, offered the practical advantage that females and males homozygous for indels could be directly sequenced. Recessive alleles on the Z chromosome of females are exposed to a greater selection, and faster selective substitution, than those recessive alleles occupying an autosomal location (Charlesworth et al. 1987). Furthermore, Z chromosomes spend two thirds of their time in males (if sex ratios are equal), which may have a higher per generation mutation rate than females (Miyata et al. 1987). Taken together, this implies that the Z chromosome may evolve more quickly than its autosomal counterparts, and makes *Tpi* a particularly interesting study locus.

Although important advances have now been made in the understanding of the molecular evolution of intronic regions, such knowledge remains limited. In particular, few studies to date have investigated all known introns of a single gene (Chamary & Hurst 2004). Despite this, it is clear that intronic evolution is complex and dynamic; patterns of tempo and mode of evolution are not always applicable to the different introns within a single gene, let alone to different genomic regions or divergent taxa. To contribute to the literature on intron evolution, we discuss intron base composition and compare *Tpi* introns to silent exonic sites.

Materials and methods

Sampling method

Samples for DNA analysis were collected, preserved, identified, the DNA extracted and dried wings retained, as previously described (Whinnett et al.

2005b). Where possible, study individuals were the same as those previously sequenced for mtDNA, and followed the replicated design of three individuals representing each of the adjacent centres of endemism. Modifications to the study taxa were as follows; nine species were removed as they were only represented by one subspecies (*Napeogenes larina*, *Napeogenes pharo*, *Ithomia lagusa*, *Hypothyris anastasia*, *Hypothyris fluonia*, *Hypothyris ninonia*, *Hypothyris* sp nov, *Hypothyris cantobrica* and *Melinaea ludovica*). Twenty-six individuals were removed as they belonged to a subspecies already represented in triplicate and an additional *Melinaea isocomma* was newly included. In addition, seven replacements were made due to the unavailability of template DNA (*Hyposcada* 1293 with 511, 1644 and 716 with 1737 and 715, 1666 with 950, *Oleria* 1304 with 286, and 1580 with 195, *Melinaea* 613 with 619).

Primer development and PCR protocols

Initial *Tpi* PCR amplifications were performed in a 25 µl volume reaction, using 2 µl *Hyposcada* or *Oleria* template DNA (specimens 715, 835, 1644, 1645 and 1674) and the following conditions: 1x PCR buffer, 1.2 mM dNTPs, 6 mM $MgCl_2$, 0.25 mM both Tpi-1 and Tpi-2 primers (Table 1), 0.025 U/µl Taq, and an amplification profile of 94°C for 2 min, followed by 35 cycles of (94°C for 45 sec, 49°C for 45 sec, 72°C for 90 sec) and a final 10 min extension at 72°C. PCR products were purified using a commercial kit, according to the manufacturer's protocol (QIAquick: Qiagen). Direct sequencing was unsuccessful, so PCR products were directly cloned into the pGEM-T Easy Vector (Promega) according to manufacturer's instructions, and plasmid preps were made using QIAprep columns (Qiagen). Fifteen µl cycle sequencing reactions, comprised of; 2 µl miniprep product, 1 µl BigDye Terminator, 1x sequencing buffer and 0.16 mM standard vector primer T7b 5′-GACTCACTATAGGGCGAATTG-3′ or SP6 5′-GTTGGATGCATAGCTTGAGTATTC-3′, were performed according to ABI recommendations. Sequencing gels were run by a commercial facility.

Primers corresponding to *Heliothis Tpi* (GenBank U23080) exons 1-5 were developed in three stages (Table 1). The following PCRs were all performed with 0.5 U/µl Taq, 0.8 mM dNTPs, a 3 min initial denaturation at 94°C, and a 34 cycle profile with a 60 sec denaturation at 94°C, 60 sec annealing, and a 60 sec extension at 72°C, with additional parameters described in Table 2. First-stage

primers TPI-5, TPI-6 and TPI-7 were designed to amplify the region corresponding to *Heliothis Tpi* exon 1 through exon 4, in 5 specimens (primer taxa 151, 707, 715, 1650, 1667), representating the four ithomiine genera: *Hyposcada, Ithomia, Hypothyris* and *Melinaea*. A second round of primers (TPI-A to TPI-F) were designed to amplify an overlapping region corresponding to *Heliothis Tpi* exon 2 or 3 through exon 5, in the 5 primer taxa. Final pairwise primers (Ali, Nab, Jess and Hav) were then designed and final PCRs were typically performed directly, using Ali with Nab, or Jess with Nab. In cases where template DNA only produced a weak band, a product was amplified using Ali with Nab then re-amplified using Jess with Nab, with both rounds of amplification reduced to just 20 cycles. PCR products were gel extracted, incubated at 70 °C for 10 mins, and incubated for 3 h at 45°C, with 1 µl GELase (Epicenter Technologies, Madison WI). Two µl of this product, 1 µl BigDye Terminator, 1 mM primer and 6 µl ddH_20 cycle sequencing reactions were performed following the ABI protocol. Cycle sequencing products were diluted in 10 µl ddH_20, purified in Centrisep columns filled with 700 µl G-50 Sephadex, and air dried. Each sample was resuspended in 10 µl Hi-Di formamide, denatured at 95 °C for 2 min and then run on an ABI 3100 Automated DNA Sequencer following the manufacturer's protocol.

Sequence analysis

Sequences were edited using SeqEd v1.0.3 and compared to *Tpi* sequences already available on Genbank. Bayesian analysis was performed on the exon sequence data of all ithomiine specimens, plus the outgroup *Helicoverpa* (Genbank Accession AY736358), using MrBayes 3.0 (Huelsenbeck & Ronquist 2001), with parameters identified by Modeltest (nst=6 and a gamma rates heterogeneity model) with 4 simultaneous chains run for 1,000,000 generations, sampling a tree every 100 generations. A consensus tree with branch support in the form of posterior probabilities was derived from the final 9,000 trees (representing the final 900,000 generations), after confirmation that likelihood values had stabilised after the first 100,000 generations.

Table 1. Name, sequence, location and origins of: the 4 final primers (Ali, Jess, Nab, Hav); and the primers used during the development of the final primers.

Primer	5'-3' Sequence (direction)	Exon location	Amplifies exons	Details in or designed from:
Tpi-1	GGTCACTCTGAAAGGAGAACCATCTT (F)	3	3-4	Beltrán et al 2002
Tpi-2	CACAACATTTGCCCAGTTGTTGCCAA (R)	4	3-4	Beltrán et al 2002
TPI-5	ACTGGAAGATGAAYGGNGAC (F)	1	1-4	*Helicoverpa* AY736358, *Spodoptera*, L39011, *Drosophila* X57576, U60848, AF025814, AF025815
TPI-6	TGAATACYACCTCCTCAGTCT (R)	4	1-4	Oleriini 715, 835, 1644, 1645, 1674
TPI-7	TCACCAATACARGCWATCACTTTC (R)	4	1-4	*Dryas iulia* AF413797, Oleriini 715, 835, 1644, 1645, 1674
TPI-A	ATCTGATCANTTAGTTGCTG (F)	3	3-5	Oleriini 715, 835, 1644, 1645, 1674
TPI-B	GAATCTGATCAATTAGTGGCTGATAA (F)	3	3-5	*Ithomia* 1650
TPI-C	GATCAAGGATGTTGGTGCAGATT (F)	3	3-5	*Hypothyris* 707
TPI-D	TTGCYGCWCAAAACTGCT (F)	2	2-5	*Melinaea* 151
TPI-E	TCACGGATCCACCGTACT (R)	5	3-5	*Heliothis* U23080, *Drosophila* X57576
TPI-F	ATACGNACNGANTGNGCNACNTC (R)	5	3-5	*Hyposcada* 715, 1667, *Helicoverpa* AY736358, *Spodoptera* L39011, *Drosophila* X57576, U60848, AF025814, AF025815
Ali	AGCAANCAGATWAATGAAATC (F)	1	1-4 or 1-5	Ithomiines 151, 707, 1650, 715, 1667
Jess	AGCAANCAGATWAATGAAATCRTYA (F)	1	1-5	Ithomiines 151, 707, 1650, 715, 1667
Nab	TAGCCAYTTACGNAGAGANGC (R)	5	1-5	Ithomiines 151, 707, 1650, 715, 1667
Hav	GTTCCAATNGCCCANACTG (R)	4	1-4	Ithomiines 151, 707, 1650, 715, 1667

Table 2. Primer pairs as used in this study, with optimised PCR parameters.

Primer combination	µl of template DNA	Specimen numbers	Reaction volume in µl	mM of each primer	$MgCl_2$ in mM	Amplification temperature in °C	Final extension at 72°C in min
TPI5 & TPI6	1.6	1667, 715, 707 & 151	10	0.50	1	49.0	15
TPI5 & TPI7	1.6	1650	10	0.75	1	49.0	15
TPI-A & TPI-E	1.6	1667, 715	10	0.50	1	49.0	15
TPI-B & TPI-F	1.6	1650	10	0.66	1	53.0	15
TPI-C & TPI-F	1.6	707	10	0.66	1	47.6	15
TPI-D & TPI-F	1.6	151	10	0.66	1	58.0	15
Ali & Nab, Jess & Nab, or Ali & Hav	2	all final amplifications	15	0.66	1.66	56.0	10

To ensure accurate comparisons, data sets for subsequent analyses only included the 84 specimens sequenced for both mtDNA and *Tpi*, excluding the morphologically identified hybrids (1358, 1541, 1958 and 2091) (Whinnett et al 2005b). The comparative mtDNA data set included just the coding COI and COII (not the intervening tRNA). PAUP version 4.0b 10 (Swofford 2000) was used to calculate the number of variable and parsimony informative sites. The Jukes-Cantor (JC) correction was chosen to enable direct comparison to values generated by Syn-SCAN (http://hivdb.stanford.edu/pages/synscan.html), JC was used to calculate corrected pairwise divergences, nucleotide composition, and transition and transversion statistics. Syn-SCAN was used to investigate synonymous and non-synonymous sites, using the weighted scoring matrix option. Syn-SCAN generates: the number of observed synonymous (S_d) and non-synonymous substitutions (N_d), the average number of potential synonymous (S) and non-synonymous (N) substitutions for the two compared sequences, the proportion of observed synonymous (p_S) and non-synonymous (p_N) substitutions, S_d/S and N_d/N, JC corrected P_S (d_S) and P_N (d_N), and the d_N/d_S ratio for all within genera comparisons. Rates of evolution between genes were compared by plotting mean within genera pairwise JC corrected sequence divergences of *Tpi* introns against d_S values for *Tpi* exon and coding mitochondrial regions.

Results

Sequence results and phylogenetic analysis

Tpi sequences were successfully obtained in 95 individuals (Table 3). Of these, four individuals (8, 154, 619 and 1541) had proved unsuccessful with the above primer combinations so were amplified with Ali & Hav. A further eight individuals failed to PCR amplify despite repeated attempts, and 19 males had unreadable chromatograms thus were assumed to be heterozygous for indels.

The phylogenetic hypothesis recovered by Bayesian analysis is shown in Fig. 1. Phylogenetic resolution was less clear than previously recovered for mtDNA (Whinnett et al. 2005b), in part due to the shorter sequence length of amplified *Tpi* exon (567 bp *Tpi* exon vs 1619 bp mtDNA). As found with the mtDNA data set, all genera formed monophyletic clades with the exception of *Hypothyris* + *Hyalyris*, and within genera *Tpi* haplotypes tended to cluster by species, with three species of *Melinaea* (with mtDNA, *Melinaea isocomma*

Table 3. Specimens sequenced from San Martin, Peru, indicating the subspecies, number and sex of specimens, and exact collection localities and co-ordinates. * represents missing data, and ** indicates that the sex of the individual is unknown.

Taxa	PERU: San Martin locality	Specimen ID, males are indicated in [ ]	Latitude	Longitude
Pseudoscada florula gracilis	Km 28 Tarapoto-Yurimaguas	1914	06°41'S	76°31'W
Pseudoscada florula gracilis	Km 42 Tarapoto-Yurimaguas	[1967]	*	*
Pseudoscada florula aureola	Km 28 Tarapoto-Yurimaguas	2086	06°41'S	76°31'W
Pseudoscada florula gracilis x aureola	Km 42 Tarapoto-Yurimaguas	[1358]	*	*
Pseudoscada florula gracilis x aureola	Km 26 Yurimaguas-Tarapoto	[2091]	*	*
Napeogenes inachia pozziana	Km 7.2 Carretera Pongo-Barranquita	261, 262, 495	06°17'S	76°13'W
Napeogenes inachia ssp nov	Km 28 Tarapoto-Yurimaguas	466, 1188, 1780	06°41'S	76°31'W
Napeogenes sylphis corena	Km 26 Yurimaguas-Tarapoto	2042, 2043, 2044	*	*
Brevioleria arzalia arzalia	Km 28 Tarapoto-Yurimaguas	1200, 1203, 1769	06°41'S	76°31'W
Brevioleria arzalia ssp nov	Km 7.2 Carretera Pongo-Barranquita	[2108]	06°17'S	76°13'W
Brevioleria aelia orolina	Km 26 Yurimaguas-Tarapoto	[2020]	*	*
Brevioleria aelia cf *pachiteae*	Serranoyacu	819, 821, [1665]	05°40'S	77°40'W
Hyalyris oulita oulita	Pte. Aguas Verdes	[853]	05°39'S	77°38'W
Hyalyris oulita outlia	Serranoyacu	811, 812, 813	05°40'S	77°40'W
Dircenna adina xanthophane	Serranoyacu	794	05°40'S	77°40'W
Dircenna adina xanthophane	Pte. Aguas Verdes	880	05°39'S	77°38'W
Dircenna adina xanthophane	La Antena	1601	06°27'S	76°17'W
Dircenna dero ssp nov	Km 7.2 Carretera Pongo-Barranquita	322, 525, 1318	06°17'S	76°13'W
Dircenna dero ssp nov	Serranoyacu	760	05°40'S	77°40'W
Hyposcada anchiala interrupta	Km 7.2 Carretera Pongo-Barranquita	511, 512, 2105	06°17'S	76°13'W
Hyposcada anchiala mendax	Serranoyacu	715, 1645, [1644]	05°40'S	77°40'W
Hyposcada anchiala mendax	La Antena	[2141]	06°27'S	76°17'W
Hyposcada zarepha flexibilis	Chumia	[197]	06°36'S	76°11'W
Hyposcada zarepha flexibilis	Shapaja	[1591]	*	*
Hyposcada zarepha ssp nov	Serranoyacu	[950], 1667, 1670	05°40'S	77°40'W
Oleria agarista agarista	Km 7.2 Carretera Pongo-Barranquita	1307	06°17'S	76°13'W
Oleria ilerdina lerida	Km 7.2 Carretera Pongo-Barranquita	[291], 1302	06°17'S	76°13'W
Oleria gunilla lota	Km 7.2 Carretera Pongo-Barranquita	286, 290, 2118	06°17'S	76°13'W
Oleria gunilla serdolis	Chumia	194, 195, 196	06°36'S	76°11'W
Oleria onega ssp nov	Shapaja	1173, 1585	*	*
Oleria onega ssp nov	Serranoyacu	835, [1674]	05°40'S	77°40'W
Oleria onega janarilla	Shapaja	1165, 1583, 1960	*	*
Oleria onega ssp nov *x janarilla*	Shapaja	1958	*	*
Ithomia agnosia agnosia	Pte. Aguas Verdes	[914]	05°39'S	77°38'W
Ithomia agnosia agnosia	Serranoyacu	1650, 1651	05°40'S	77°40'W
Ithomia agnosia agnosia	Chumia	1450	06°36'S	76°11'W
Ithomia agnosia agnosia	Shapaja	1590	*	*
Ithomia salapia aquina	Km 30 Yurimaguas-Tarapoto	71, 72, 74	*	*
Ithomia salapia aquina	Chumia	203, 204, 1445	06°36'S	76°11'W
Ithomia salapia derasa	Serranoyacu	733	05°40'S	77°40'W
Hypothyris euclea ssp nov	Pte. Aguas Verdes	868, 894	05°39'S	77°38'W
Hypothyris euclea ssp nov	Shapaja	1098, 1099, 1103	*	*
Hypothyris mansuetus meterus	Serranoyacu	707	05°40'S	77°40'W
Hypothyris mansuetus meterus	Pte. Aguas Verdes	874	05°39'S	77°38'W
Melinaea menophilus ssp nov	*	8, 9, 619	06°27'S	76°20'W
Melinaea menophilus ssp nov	Serranoyacu	[683]	05°40'S	77°40'W
Melinaea menophilus hicetas x ssp nov	Shapaja	1541	*	*
Melinaea isocomma simulator	Serranoyacu	688, 689, 690, 1716	05°40'S	77°40'W
Melinaea marsaeus mothone	Serranoyacu	684, 686, [687]	05°40'S	77°40'W
Melinaea marsaeus phasiana	Chumia	151, 154	06°36'S	76°11'W
Melinaea marsaeus rileyi	Km 7.2 Carretera Pongo-Barranquita	233	06°17'S	76°13'W
Melinaea satevis tarapotensis	Shapaja	1082**	*	*
Melinaea satevis cydon	Km 7.2 Carretera Pongo-Barranquita	1260	06°17'S	76°13'W

increased this to four species of *Melinaea*) and the *Oleria onega* proving to be exceptions. In contrast to mtDNA, which recovered *Brevioleria aelia* as paraphyletic with respect to *Brevioleria arzalia*, the *Tpi* exon data recovered both *Brevioleria aelia* subspecies as basal to the *Brevioleria* clade. Many subspecies clusters are recovered with the *Tpi* exon, but it typically did not identify such strong differentiations as found using the mtDNA, and there was one case (*Hyposcada zarepha*) of two subspecies clearly distinguished with mtDNA not forming distinct clades using *Tpi.* Finally, subspecies pairs which had not attained complete distinction at mtDNA were similarly recovered using *Tpi*: *Napeogenes inachia*, *Pseudoscada florula*, *Brevioleria arzalia*, *Oleria gunilla*, and *Ithomia salapia.*

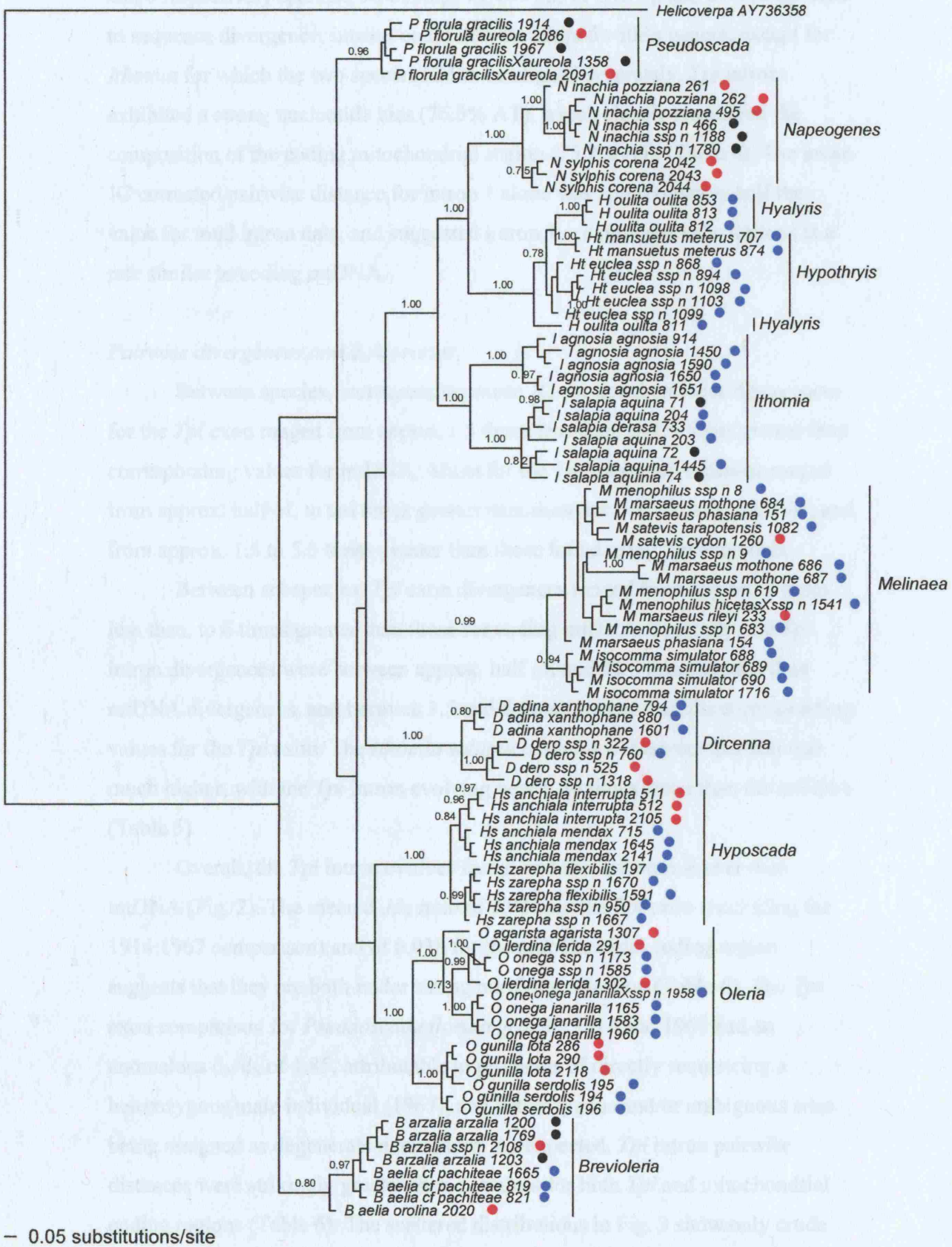
Tpi exon and Tpi intron

There was no length variation in the exon regions studied. The full exon alignment (corresponding to the region between primers TPI-5 and TPI-E) corresponds to the most 3' 63 bp, complete 112, 110 and 207 bp, and the most 5' 79 bp of exons 1-5 of *Heliothis Tpi* (Genbank U23080). Of the 571 bp total exon, 26.5% of the sites were variable, a value comparable to that for coding mtDNA (29.0%). Of the *Tpi* exons for which complete reads were available, exon 4 evolved slightly more rapidly (exons 2, 3 and 4 had 25.0, 26.4 and 32.4% variable sites, respectively). As expected, third positions of the codons harboured the highest proportion of variable sites with positions 1, 2 and 3 respectively being 7.9%, 6.3% and 64.7% variable. *Tpi* exons were composed of a relatively equal base composition (54.6% AT) (Table 4), similar to previously observed compositions of coding nuclear regions in ithomiine butterflies; *Tektin* 59.9%, *wingless* 46.2% and *EF 1-α* 49.5% AT (Whinnett et al. 2005a).

Table 4. Nucleotide composition of the aligned coding mitochondrial (1555 bp), *Tpi* exon (567 bp) and *Tpi* intron (418-573 bp) regions, for the 84 ithomiine specimens sequenced for both mtDNA and *Tpi*. Morphologically identified hybrids (1358, 1541, 1958 and 2091) are excluded.

	coding mitochondrial	*Tpi* exon						*Tpi* intron
		1	2	3	4	5	1-5	
A (%)	33.4	26.5	29.6	26.3	29.7	26.3	28.7	34.3
C (%)	13.3	24.8	18.7	15.9	18.7	27.0	18.3	11.4
G (%)	11.5	30.4	25.3	27.9	27.5	23.4	27.1	12.0
T (%)	41.8	18.4	26.4	29.9	24.0	23.4	25.9	42.2

Figure 1. Phylogenetic hypothesis based on *Tpi* exon data for *Dircenna* (D), *Hyalyris* (H), *Brevioleria* (B), *Hyposcada* (Hs), *Hypothyris* (Ht), *Ithomia* (I), *Melinaea* (M), *Napeogenes* (N), *Oleria* (O) and *Pseudoscada* (P). Tree topology is consensus of the last 9000 trees inferred using Bayesian methods. Hybrid individuals (1358, 1541, 1958 and 2091) are identified using morphological characters. ● indicates those specimens from the Huallaga zone, ●the Ucayali zone, and ●the suture zone.



Complete intron 1-4 sequences ranged from 418-419 bp (*Hypothyris mansuetus*) to 573 bp (*Hyposcada* 1645). The point of intron insertions was conserved in all ithomiines. All introns exhibited length variations; introns 1, 2, 3 and 4 respectively spanned 96-213 bp, 94-179 bp, 81-215 bp and 64-110 bp. Due to sequence divergence, introns could only be aligned within genera, except for *Ithomia* for which the two species had to be aligned separately. *Tpi* introns exhibited a strong nucleotide bias (76.5% AT), which closely resembled the composition of the coding mitochondrial region (75.2% AT) (Table 4). The mean JC corrected pairwise distance for intron 1 alone was approximately half the value for total intron data, and suggested intron 1 accumulates substitutions at a rate similar to coding mtDNA.

Pairwise divergences and d_N/d_S ratios

Between species, corresponding mean JC corrected pairwise divergences for the *Tpi* exon ranged from approx 1.5 times less than, to 1.5 times greater than corresponding values for mtDNA, whilst for the *Tpi* intron, divergences ranged from approx. half of, to ten times greater than those recovered using mtDNA, and from approx. 1.5 to 5.5 times greater than those found using *Tpi* exon data.

Between subspecies, *Tpi* exon divergences ranged from approx. 6 times less than, to 6 times greater than those for coding mtDNA. Typically, the *Tpi* intron divergences were between approx. half of, and 12.5 times greater than mtDNA divergences, and between 1.5 and 8 times greater than the corresponding values for the *Tpi* exon. The *Ithomia salapia aquina* x *derasa* comparison was much higher, with the *Tpi* intron evolving nearly 30 times faster than the mtDNA (Table 5).

Overall, the *Tpi* intron evolves faster, and the *Tpi* exon slower than mtDNA (Fig. 2). The mean d_N/d_S ratio of 0.041 for the *Tpi* exon (excluding the 1914:1967 comparison) and of 0.038 for the mitochondrial coding region suggests that they are both under strong purifying selection (Table 6). The *Tpi* exon comparison for *Pseudoscada florula gracilis* 1914 and 1967 had an anomalous d_N/d_S of 1.85, attributable to problems of directly sequencing a heterozygous male individual (1967), and heterozygous and/or ambiguous sites being assigned as degenerate nucleotides. As expected, *Tpi* intron pairwise distances were strikingly greater than d_N values for both *Tpi* and mitochondrial coding regions (Table 6). The scattered distributions in Fig. 3 show only crude

correlations between *Tpi* intron pairwise distances and d_S values for both *Tpi* exon and mitochondrial coding regions. However, the mean d_S values of 0.047 for *Tpi* exon, and 0.084 for mitochondrial coding regions indicate that silent sites are accumulating slightly more rapidly in the *Tpi* exon, and more than twice as rapidly in mitochondrial coding regions, than in the total intronic sequence (mean JC corrected pairwise distance 0.038), and intron 2-4 sequence (mean JC corrected pairwise distance 0.042) (Table 6).

Discussion

Phylogenetic relationships

Both the mtDNA and *Tpi* loci recovered similar relationships between the taxa. For example, the *Oleria onega* (individual 1958) is a putative hybrid between the morphologically divergent *Oleria onega janarilla* and *Oleria onega* ssp nov. This individual is recovered in the *O. onega janarilla* cluster, which is interestingly markedly distinct from the *O. onega* ssp nov cluster, using both mtDNA and *Tpi* sequence data. This finding mirrors a RAPD analysis by Gallusser and colleagues (2004) which placed putative *Oleria onega agarista* X *Oleria onega* ssp nov hybrids as closer to *O. onega agarista* than *O. onega* ssp nov. Gallusser et al. (2004) referred to the subspecies now recognised as being *O. onega agarista* as *O. onega janarilla* in their paper). It would be interesting to perform a more detailed study on these *Oleria* subspecies, in an attempt to elucidate whether they diverged in allopatry and have since come into secondary contact, or whether they arose in-situ through parapatric diversification.

In stark contrast to the *Oleria onega* subspecies are *Melinaea marsaeus*, *Melinaea satevis* and *Melinaea menophilus*. These three species of *Melinaea* are indistinguishable using both mtDNA and *Tpi* loci, despite the fact that these loci are amongst the most rapidly evolving markers currently available for Lepidoptera. Proponents of DNA barcoding have suggested a 3% threshold at the COI mitochondrial locus for Lepidoptera species (Hebert et al 2003). Not only would a DNA taxonomy based on COI have failed to recognise these *Melinaea* species as distinct, but a DNA taxonomy based on a nuclear locus such as *Tpi* (as proposed as a possible supplement to strengthen mitochondrial DNA taxonomy) would have also failed. Although DNA taxonomy has now been shown to successfully distinguish between a number of taxa, cases such as this *Melinaea* example present an important problem which should be addressed by DNA barcoders.

Table 5. Mean Jukes Cantor corrected % pairwise divergences of the aligned coding mitochondrial (1555 bp), *Tpi* exon (567 bp) and *Tpi* intron (418-573 bp) regions, for the 84 ithomiine specimens sequenced for both mtDNA and *Tpi* (morphologically identified hybrids are excluded). A value is not available for the *Ithomia agnosia* and *I. salapia Tpi* intron comparison, as it was not possible to align the sequences due to high divergence. Mitochondrial DNA results here differ from Whinnett et al. (2005b) due to the different model of sequence evolution, exclusion of tRNA, and only a subset of the total mtDNA sequences being analysed here.

BETWEEN GENERA	Coding mitoch-ondrial	*Tpi* exon
Pseudoscada x Napeogenes	7.95	8.77
Pseudoscada x Brevioleria	7.85	5.28
Pseudoscada x Hyalyris	7.95	9.04
Pseudoscada x Dircenna	7.47	7.27
Pseudoscada x Hyposcada	7.88	6.60
Pseudoscada x Oleria	8.82	5.56
Pseudoscada x Ithomia	8.08	7.78
Pseudoscada x Hypothyris	8.92	8.59
Pseudoscada x Melinaea	9.52	9.27
Napeogenes x Brevioleria	8.58	9.00
Napeogenes x Hyalyris	7.83	6.39
Napeogenes x Dircenna	7.94	10.66
Napeogenes x Hyposcada	8.03	10.04
Napeogenes x Oleria	9.13	9.28
Napeogenes x Ithomia	8.23	7.31
Napeogenes x Hypothyris	8.46	5.96
Napeogenes x Melinaea	9.24	10.46
Brevioleria x Hyalyris	7.96	9.78
Brevioleria x Dircenna	7.97	8.20
Brevioleria x Hyposcada	8.55	8.06
Brevioleria x Oleria	9.22	5.94
Brevioleria x Ithomia	8.52	8.86
Brevioleria x Hypothyris	9.10	9.29
Brevioleria x Melinaea	9.91	10.76
Hyalyris x Dircenna	7.30	10.33
Hyalyris x Hyposcada	7.80	11.13
Hyalyris x Oleria	9.24	9.21
Hyalyris x Ithomia	7.70	7.20
Hyalyris x Hypothyris	4.60	1.53
Hyalyris x Melinaea	9.60	11.06
Dircenna x Hyposcada	7.26	8.63
Dircenna x Oleria	9.02	7.32
Dircenna x Ithomia	7.79	9.99
Dircenna x Hypothyris	8.33	9.71
Dircenna x Melinaea	9.49	11.62
Hyposcada x Oleria	8.28	6.68
Hyposcada x Ithomia	8.41	9.59
Hyposcada x Hypothyris	8.65	10.72
Hyposcada x Melinaea	10.07	12.15
Oleria x Ithomia	9.59	9.05
Oleria x Hypothyris	9.72	8.98
Oleria x Melinaea	9.87	10.44
Ithomia x Hypothyris	8.62	7.22
Ithomia x Melinaea	10.00	10.63
Hypothyris x Melinaea	10.52	11.45

BETWEEN SPECIES	Coding mitoch-ondrial	*Tpi* exon	*Tpi* intron
Napeogenes inachia x sylphis	3.52	1.60	4.01
Brevioleria arzalia x aelia	0.88	0.31	1.34
Dircenna adina x dero	1.27	1.72	4.53
Hyposcada anchiala x zarepha	3.19	1.61	9.07
Oleria agarista x ilerdina	7.70	1.68	5.76
Oleria agarista x gunilla	8.48	2.84	12.4
Oleria agarista x onega	6.93	1.68	4.99
Oleria ilerdina x gunilla	7.66	2.73	11.32
Oleria ilerdina x onega	5.48	0.89	3.01
Oleria gunilla x onega	7.75	2.83	10.50
Ithomia agnosia x salapia	6.22	5.02	-
Hypothyris euclea x mansuetus	3.11	1.20	5.00
Melinaea menophilus x isocomma	0.20	0.99	1.99
Melinaea menophilus x marsaeus	0.42	1.18	3.15
Melinaea menophilus x satevis	0.90	0.53	2.87
Melinaea isocomma x marsaeus	0.35	2.04	3.05
Melinaea isocomma x satevis	0.80	1.51	3.32
Melinaea marsaeus x satevis	0.93	1.38	4.15
BETWEEN SUBSPECIES			
Pseudoscada florula gracilis x aureola	0.81	0.81	1.28
Napeogenes inachia pozziana x ssp nov	0.25	0.38	3.16
Brevioleria arzalia arzalia x ssp nov	0.66	0	0.40
Brevioleria aelia cf *pachiteae x orolina*	1.45	0.81	2.88
Hyposcada anchiala interrupta x mendax	0.59	0.34	2.60
Hyposcada zarepha flexibilis x ssp nov	1.86	0	1.44
Oleria gunilla lota x serdolis	0.51	0	0.78
Oleria onega janarilla x ssp nov	5.63	1.38	3.46
Ithomia salapia aquina x derasa	0.07	0.84	1.97
Melinaea marsaeus mothone x phasiana	0.55	2.09	4.66
Melinaea marsaeus mothone x rileyi	0.53	2.26	4.55
Melinaea marsaeus phasiana x rileyi	0.31	0.72	2.61
Melinaea satevis tarapotensis x cydon	1.59	0.69	3.64

Figure 2. JC corrected *Tpi* pairwise distances, plotted as a function of the corresponding JC corrected coding mitochondrial pairwise distances, for all possible within-genera combinations of the 84 ithomiine specimens sequenced for both mtDNA and *Tpi* (morphologically identified hybrids are excluded). ■ represents intron data, and ○ represents exon data. Lines of best fit were created in Microsoft Excel.

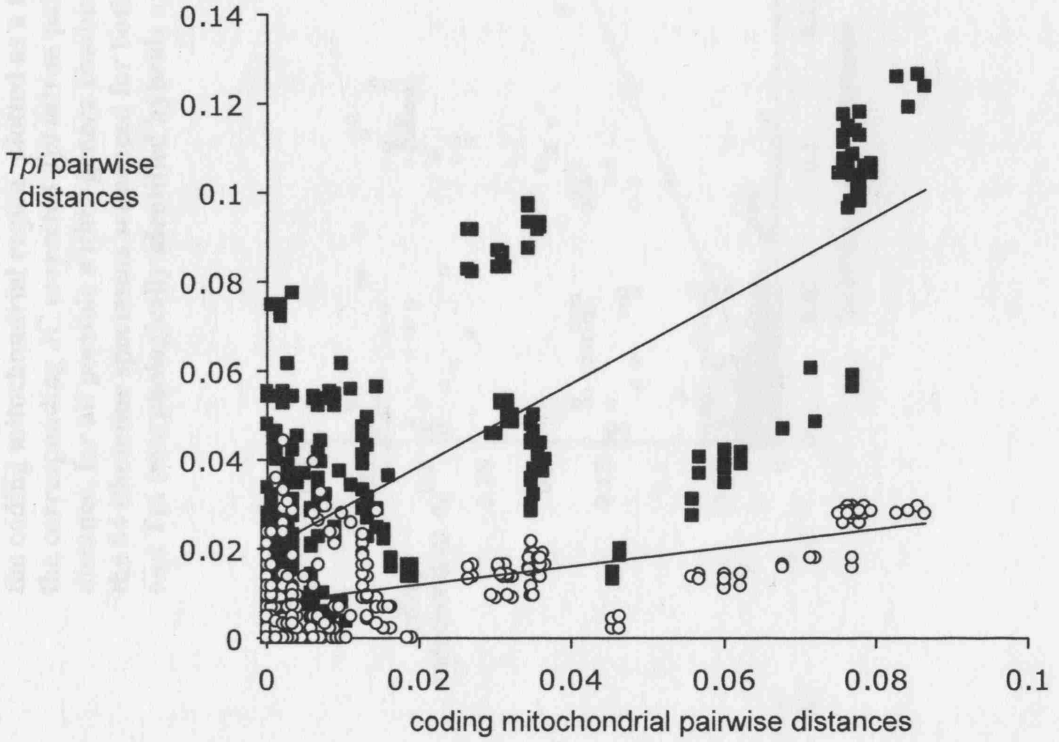


Table 6. Mean within genera JC corrected % pairwise divergences for coding mitochondrial, *Tpi* exon and *Tpi* intron regions, for the 84 ithomiine specimens sequenced for both mtDNA and *Tpi* (morphologically identified hybrids are excluded). d_N:d_S statistics are presented for coding mitochondrial and *Tpi* exon regions, for the same 84 ithomiine specimens. (d_N:d_S values are not available for the *Tpi* intron as synonymous vs nonsynonymous substitutions are not a relevant to non-coding regions.) * indicates that the 1914:1967 comparison was removed (due to an anomalous d_N/d_S of 1.85).

	Tpi exon	*Tpi* intron	coding mitochondrial
JC % pairwise divergence (minimum/maximum)	1.2 (0/4.4)	3.8 (0/12.6)	2.0 (0/8.6)
mean d_S (minimum/maximum)	0.047 (0/0.19)	-	0.084 (0/0.36)
mean d_N (minimum/maximum)	0.001 (0/0.01)	-	0.003 (0/0.03)
mean d_N/d_S (minimum/maximum)	0.041* (0/0.63*)	-	0.038 (0/0.43)

Figure 3a. JC corrected synonymous substitutions (d_S) for *Tpi* exons 1-5, plotted as a function of the corresponding JC corrected *Tpi* intron pairwise distance, for all possible within-genera combinations of the 84 ithomiine specimens sequenced for both mtDNA and Tpi (morphologically identified hybrids are excluded).

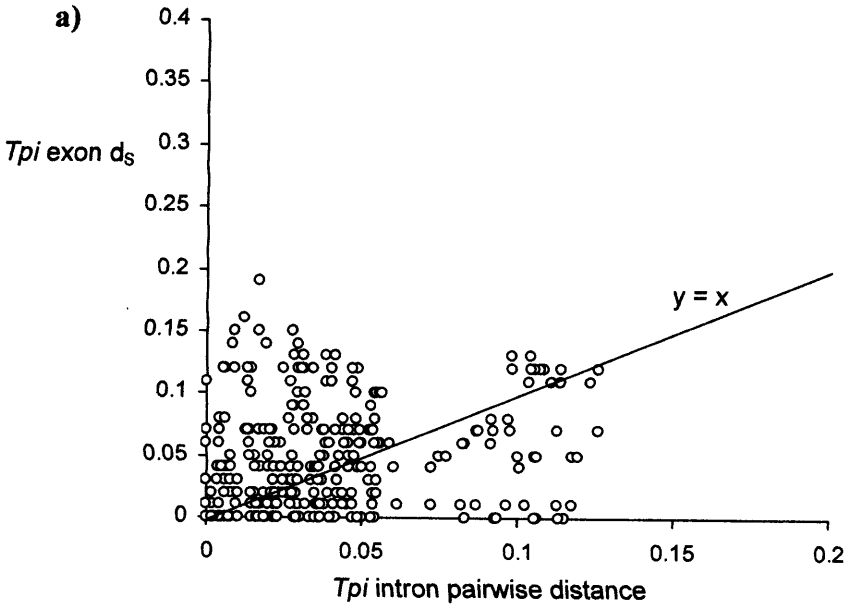
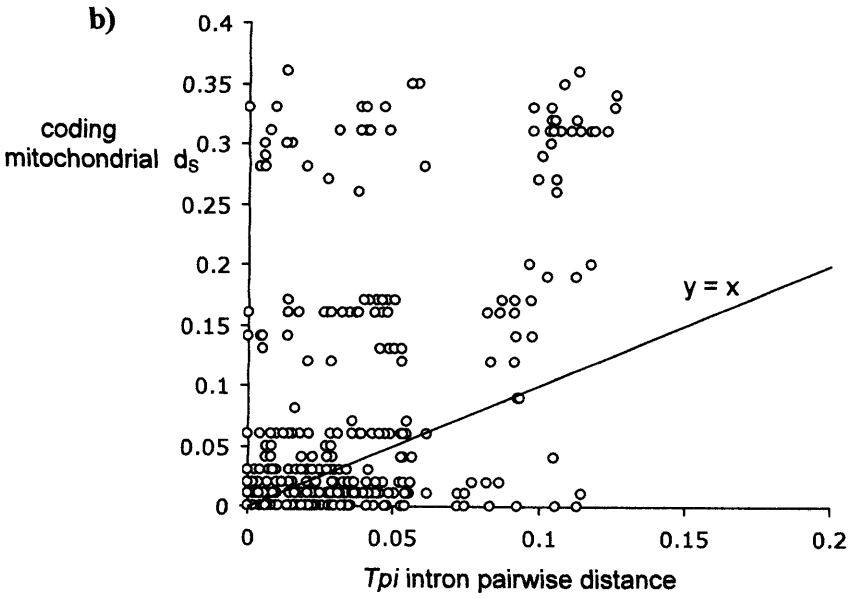


Figure 3b. JC corrected synonymous substitutions (d_S) for the coding mitochondrial region, plotted as a function of the corresponding JC corrected *Tpi* intron pairwise distance, for all possible within-genera combinations of the 84 ithomiine specimens sequenced for both mtDNA and Tpi (morphologically identified hybrids are excluded).



MtDNA and Tpi divergences

Similar to previous findings using mtDNA (Whinnett et al. 2005b), we recovered a marked variability in pairwise divergences both for species within genera, and for subspecies within species, comparisons at both the *Tpi* exon and the *Tpi* intron. Consistent with molecular clock theory, the variable *Tpi* divergences identified here appear to provide additional evidence to reject a concordant diversification of ithomiines. It therefore appears that based on evidence from two independent loci we can refute the hypothesis that simple vicariance, such as that implicated by the Pleistocene Refugia hypothesis, drove a simultaneous diversification of all ithomiine taxa studied here.

Under the assumptions of near-neutral evolution, we might expect different genomic regions within individuals to evolve somewhat in consort. Indeed, comparison of the data revealed a good correlation between divergences at mtDNA and *Tpi*, especially between mtDNA and *Tpi* exon divergences. However, evolution at mtDNA and *Tpi* was not always tightly correlated, for example, subspecies of *Ithomia salapia* were much more divergent, and subspecies of *Oleria onega* and *Brevioleria arzalia* more similar, at both the *Tpi* exon and intron than for mtDNA. This could be explained by, for example; variation in the rates of molecular evolution among distinct genomic regions due to factors such as differential functional constraints; and variation among distinct lineages due to features peculiar to the biology of each group, such as generation time, or introgressive hybridisation.

Hybridisation

Hybridisation is an important evolutionary mechanism, with maybe 12- 26% of species of butterflies known to hybridise. Even if just a few of these hybrids are fertile, introgression is possible and maybe even likely (Mallet 2005). Although morphological intermediacy likely underestimates the true number of hybrids, for example, it misses those cases in which wing pattern is determined by a few loci or supergenes, we determined from morphology that many of the subspecies from our study site were indeed capable of hybridisation. Our previous mtDNA results revealed that hybridisation was occurring even between subspecies which otherwise

retain very marked distinction, for example between *Oleria onega* subspecies, despite a 10.19% mtDNA divergence.

Introgression can affect the mitochondrial and nuclear genomes differently. Haldane (1922) observed that when hybridisation occurs, the heterogametic sex suffers a greater sterility and inviability than the homogametic sex. In addition, there is the correlated phenomenon known as the 'X effect', in which genes of importance to hybrid sterility and inviability are often X-linked (Coyne & Orr 1989). Deleterious, X-linked recessive alleles are always exposed, and therefore have a greater effect, in the heterogametic sex, whilst they can be 'hidden' by non-deleterious dominant alleles in the homogametic sex. Based on this, and that females are the heterogametic sex in Lepidoptera, we might expect a more restricted introgression (and thus stronger correlation to taxon divergences) of mtDNA, than of nuclear regions that would generally be more able to introgress though fertile, male hybrids (Tegelstrom & Gelter 1990). We initially predicted that *Tpi* might help identify hybridisation events undetected by morphology or with mtDNA but we did not identify any clear signatures of hybridisation from the *Tpi* data. Nuclear regions themselves differ in their susceptibility to infiltrate (Mallet 2005) and *Tpi* has indeed been reported to be strongly associated with female sterility in interracial and interspecific *Heliconius* hybrids (Jiggins et al. 2001, Naisbit et al. 2002). Therefore, perhaps different nuclear loci which are completely unlinked from any barrier to gene exchange (for example, the *Mpi* and *Ci* loci used by Bull (2003) to study *Heliconius*) will introgress more readily than *Tpi*, and therefore be more effective for identifying hybridisation events.

The molecular evolution of Tpi

In this section, we focus on base composition and substitution rates in *Tpi*, to improve our understanding of *Tpi* intron evolution, and the evolution of introns in general. The mean base compositions of the *Tpi* exon and *Tpi* intron observed here were 54.6 and 76.5% A+T. This finding is in contrast to previous reports of a similar base composition in introns and exons in the mouse and rat (Hughes & Yeager 1997) and in *Drosophila* (Akashi et al. 1998). However, ithomiine *Tpi* introns are long compared to typical introns in *Drosophila* (Adams et al. 2000) and our findings are

more consistent with those of DaLage and colleagues (1996) who report that long introns in drosophilids were A+T rich, even though the base composition of short introns were similar to that of flanking exons. The ithomiine A+T composition also mirrors that reported by Winnard and colleagues (2002) for teleosts (48% vs 60% A+T in exon vs intron), even though the same genes in mammals exhibited similar base compositions between the exon and intron (45% and 47% A+T). Several explanations have been offered to account for base composition, including selection, tRNA availability and mutational bias. Bernardi (2000) suggested that the warm body temperature of mammals selected for a decrease in A+T nucleotides, which have a lower melting temperature than C+G. Winnard et al (2002) investigated the hypothesis that intron composition was an adaptation to thermal barriers in teleosts, but found no correlation with habitat temperature range. It would be of interest to investigate whether the high A+T recovered here is a general characteristic of Lepidoptera introns.

It has been assumed that non-coding regions and synonymous substitutions are not subject to natural selection and are therefore putatively neutral. A number of authors report findings consistent with this, for example, Hughes and Yeager (1997) found that the average rate of nucleotide substitutions in introns was very similar to the rate of synonymous substitutions in rodents. However, our results indicate that the substitution rate in *Tpi* intron was slightly less than the synonymous substitution rate of *Tpi* exon and approximately half the synonymous rate in mtDNA. This mirrors a previous study on *Heliconius* butterflies (Beltrán et al. 2002) which suggested that the neutral substitution rate was faster in mitochondria than in *Tpi* introns. Indeed, evidence is mounting that putatively neutral genomic regions may often evolve non-neutrally. For example, Hellmann and colleagues (2003) reported that as many as 40% of four-fold degenerate sites are under selection. Halligan and colleagues (2004) reported variable effects of selection on base pairs 3 to 6 at the 5′ end of introns (weaker in *D. melanogaster* than *D. simulans*) and absolute constraint at splice sites, but that in other regions of the intron, the substitution rate was higher than for synonymous sites. A third paper (Subramaniam & Kumar 2003) reported a 30-60% higher rate of synonymous substitutions in exons than in longer intergenic fragments of human and chimp.

There is growing evidence that introns have important functional roles, for example, as splicing control regions or other regulatory elements (Leicht et al. 1995, Clark et al. 1996, Majewski & Ott 2002, Hare & Palumbi 2003, Sorek & Ast 2003). As strong evidence for this, we found that the mean JC corrected pairwise distance for intron 1 was approximately half the value for the total intronic data. Although this contrasts the findings of Levy and colleagues (2001) claiming that, if anything, first introns evolve faster, our finding is consistent with reports that intron-associated regulatory elements are often located in the first intron and Chamary and Hurst's (2004) recent detailed study on mutation rates, which also found that first introns evolve more slowly in rodents.

Fast-evolving nuclear regions will almost certainly be increasingly employed as molecular markers. Introns are likely to be amongst choice regions as conserved primer binding sites located in exons can be used to amplify the intervening, hypervariable segment. However, the molecular evolution of introns is quite poorly understood. Therefore, further work into the molecular evolution of introns which are used for, or are candidates for, phylogenetic and phylogeographic work would be very useful.

References

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., Chen, L. X., Brandon, R. C., Rogers, Y. H. C., Blazej, R. G., Champe, M., Pfeiffer, B. D., Wan, K. H., Doyle, C., Baxter, E. G., Helt, G., Nelson, C. R., Miklos, G. L. G., Abril, J. F., Agbayani, A., An, H. J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R. M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E. M., Beeson, K. Y., Benos, P. V., Berman, B. P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M. R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K. C., Busam, D. A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J. M., Cawley, S., Dahlke, C., Davenport, L. B., Davies, A., de Pablos, B., Delcher, A., Deng, Z. M., Mays, A. D., Dew, I., Dietz, S. M., Dodson, K., Doup, L. E., Downes, M., Dugan-Rocha, S., Dunkov, B. C., Dunn, P., Durbin, K. J., Evangelista, C. C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A. E., Garg, N. S., Gelbart, W. M., Glasser, K., Glodek, A., Gong, F. C., Gorrell, J. H., Gu, Z. P., Guan, P., Harris, M., Harris, N. L., Harvey, D., Heiman, T. J., Hernandez, J. R., Houck, J., Hostin, D., Houston, D. A., Howland, T. J., Wei, M. H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G. H., Ke, Z. X., Kennison, J. A., Ketchum, K. A., Kimmel, B. E., Kodira, C. D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z. W., Lasko, P., Lei, Y. D., Levitsky, A. A., Li, J. Y., Li, Z. Y., Liang, Y., Lin, X. Y., Liu, X. J., Mattei, B., McIntosh, T. C., McLeod, M. P., McPherson, D., Merkulov, G., Milshina, N. V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S. M., Moy, M., Murphy, B., Murphy, L., Muzny, D. M., Nelson, D. L., Nelson, D. R., Nelson, K. A., Nixon, K., Nusskern, D. R., Pacleb, J. M., Palazzolo, M., Pittman, G. S., Pan, S., Pollard, J., Puri, V., Reese, M. G., Reinert, K., Remington, K., Saunders, R. D. C., Scheeler, F., Shen, H., Shue, B. C., Siden-Kiamos, I., Simpson, M., Skupski, M. P., Smith, T., Spier, E., Spradling, A. C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A. H. H., Wang, X., Wang, Z. Y., Wassarman, D. A., Weinstock, G. M., Weissenbach, J., Williams, S. M., Woodage, T., Worley, K. C., Wu, D., Yang, S., Yao, Q. A., Ye, J., Yeh, R. F., Zaveri, J. S., Zhan, M., Zhang, G. G., Zhao, Q., Zheng, L. S., Zheng, X. Q. H., Zhong, F. N., Zhong, W. Y., Zhou, X. J., Zhu, S. P., Zhu, X. H., Smith, H. O., Gibbs, R. A., Myers, E. W., Rubin, G. M. & Venter, J. C. (2000) The genome sequence of *Drosophila melanogaster*. *Science*, **287**: 2185-2195.

Akashi, H., Kliman, R. M. & Eyre-Walker, A. (1998) Mutation pressure, natural selection, and the evolution of base composition in Drosophila. *Genetica*, **102-103**: 49-60.

Ballard, J. W. O. & Whitlock, M. C. (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**: 729-744.

Beltrán, M. S., Jiggins, C. D., Bull, V., Linares, M., Mallet, J., McMillan, W. O. & Bermingham, E. (2002) Phylogenetic discordance at the species boundary: comparative gene genealogies among rapidly radiating *Heliconius* butterflies. *Molecular Biology and Evolution*, **19**: 2176-2190.

Bensasson, D., Zhang, D.-X., Hartl, D. L. & Hewitt, G. M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*, **16**: 314-321.

Bernardi, G. (2000) Isochores and the evolutionary genomics of vertebrates. *Gene*, **241**: 3-17.

Brower, A. V. Z., Freitas, A. V. L., Lee, M.-M., Silva Brandão, K. L., Whinnett, A. & Willmott, K. R. (2006) Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions. *Systematic Entomology*, **31**: 288-301.

Bull, V. J. (2003) Genealogy and speciation in *Heliconius* butterflies. PhD thesis, University of London.

Chamary, J. V. & Hurst, L. D. (2004) Similar rates but different modes of sequence evolution in introns and at exonic silent sites in rodents: Evidence for selectively driven codong usage. *Molecular Biology and Evolution*, **21**: 1014-1023.

Charlesworth, B., Coyne, J. A. & Barton, N. H. (1987) The relative rates of evolution of sex chromosomes and autosomes. *American Naturalist*, **130**: 113-146.

Clark, A. G., Leicht, B. G. & Muse, S. V. (1996) Length variation and secondary structure of introns in the Mlc1 gene in six species of *Drosophila*. *Molecular Biology and Evolution*, **13**: 471-482.

Coyne, J. A. & Orr, H. A. (1989) Two rules of speciation. In *Speciation and its Consequences* (ed. D. Otte & J. A. Endler), pp. 180-207. Sunderland, Massachusetts: Sinauer Associates.

DaLage, J. L., Wegnez, M. & Cariou, M. L. (1996) Distribution and evolution of introns in *Drosophila* amylase genes. *Journal of Molecular Evolution*, **43**: 334-347.

Edwards, S. V. & Beerli, P. (2000) Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**: 1839-1854.

Gallusser, S., Guadagnuolo, R. & Rahier, M. (2004) Genetic (RAPD) diversity between *Oleria onega agarista* and *Oleria onega* ssp. (Ithomiinae, Nymphalidae, Lepidoptera) in north-eastern Peru. *Genetica*, **121**: 65-74.

Haldane, J. B. S. (1922) Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, **12**: 101-109.

Halligan, D. L., Eyre-Walker, A., Andolfatto, P. & Keightley, P. D. (2004) Patterns of evolutionary constraints in intronic and intergenic DNA. *Genome Research*, **14**: 273-279.

Hare, M. P. & Palumbi, S. R. (2003) High intron sequence conservation across three mammalian orders suggests functional constraints. *Molecular Biology and Evolution*, **20**: 969-978.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **270**: 313-321.

Hellmann, I., Zollner, S., Enard, W., Ebersberger, I., Nickel, B. & Paabo, S. (2003) Selection on human genes as revealed by comparisons to chimpanzee cDNA. *Genome Research*, **13**: 831-837.

Hudson, R. R. & Turelli, M. (2003) Stochasticity overrules the "three-times rule": genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, **57**: 182-190.

Huelsenbeck, J. P. & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**: 754-755.

Hughes, A. L. & Yeager , M. (1997) Comparative evolutionary rates of introns and exons in murine. *Journal of Molecular Evolution*, **45**: 125-130.

Hurst, G. D. D. & Jiggins, F. M. (2005) Mitochondrial DNA as a marker in population, phylogeographic, and phylogenetic studies: the effects of inherited symbionts appraised. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **272**: 1525-1534.

Jiggins, C. D., Linares, M., Naisbit, R. E., Salazar, C., Yang, Z. H. & Mallet, J. (2001) Sex-linked hybrid sterility in a butterfly. *Evolution*, **55**: 1631-1638.

Jiggins, C. D., Mavarez, J., Beltrán, M., McMillan, O., Johnston, J. S. & Bermingham, E. (2005) A genetic linkage map of the mimetic butterfly, *Heliconius melpomene*. *Genetics*, **171**: 557-570.

Leicht, B. G., Muse, S. V., Hanczyc, M. & Clark, A. G. (1995) Constraints on intron evolution in the gene encoding the myosin alkali light-chain in *Drosophila*. *Genetics*, **139**: 299-308.

Levy, S., Hannenhalli, S. & Workman, C. (2001) Enrichment of regulatory signals in conserved non-coding genomic sequence. *Bioinformatics*, **17**: 871-877.

Logsden, J. M., Tyshenko, M. G., Dixon, C., Javari, J. D., Walker, V. K. & Palmer, J. D. (1995) Seven newly discovered intron positions in the triose phosphate isomerase gene: evidence for the introns-late theory. *Proceedings of the National Academy of Sciences, USA*, **92**: 8507-8511.

Majewski, J. & Ott, J. (2002) Distribution and characterization of regulatory elements in the human genome. *Genome Research*, **12**: 1827-1836.

Mallarino, R., Bermingham, E., Willmott, K. R., Whinnett, A. & Jiggins, C. D. (2005) Molecular systematics of the butterfly genus *Ithomia* (Lepidoptera: Ithomiinae): a composite phylogenetic hypothesis based on seven genes. *Molecular Phylogenetics and Evolution*, **34**: 625-644.

Mallet, J. (2005) Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, **20**: 229-237.

Miyata, T., Hayashida, H., Kuma, K. & Yasunaga, T. (1987) Male-driven molecular evolution demonstrated by different rates of silent substitutions between autosome-linked and sex chromosome-linked genes. *Proceedings of The Japan Academy Series B- Physical and Biological Sciences*, **63**: 327-331.

Naisbit, R. E., Jiggins, C. D., Linares, M., Salazar, C. & Mallet, J. (2002) Hybrid sterility, Haldane's Rule and speciation in *Heliconius cydno* and *H. melpomene*. Genetics, **161**: 1517-1526.

Pamilo, P. & Nei, M. (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution*, **5**: 568-583.

Sorek, R. & Ast, G. (2003) Intronic sequences flanking alternatively spliced exons are conserved between human and mouse. *Genome Research*, **13**: 1631-1637.

Subramaniam, S. & Kumar, S. (2003) Neutral substitutions occur at a faster rate in exons than in noncoding DNA in primate genomes. *Genome Research*, **13**: 838-844.

Swofford, D. L. (2000) *PAUP*: Phylogenetic Analysis Using Parsiomony (*and Other Methods)*. Sunderland, Massachusetts: Sinauer Associates.

Tegelstrom, H. & Gelter, H. (1990) Haldane's rule and sex-biased gene flow between two hybridizing flycatcher species *Fidecula albicollis* and *F. hypoleuca*, Aves: Muscicapidae). *Evolution*, **44**: 2012-2021.

Whinnett, A., Brower, A. V. Z., Lee, M.-M., Willmott, K. R. & Mallet, J. (2005a) The phylogenetic utility of *Tektin*, a novel region for inferring systematic relationships amongst Lepidoptera. *Annals of the Entomological Society of America*, **98**: 873-886.

Whinnett, A., Zimmermann, M., Willmott, K. R., Herrera, N., Mallarino, R., Simpson, F., Joron, M., Lamas, G. & Mallet, J. (2005b) Strikingly variable divergence times inferred across an Amazonian butterfly 'suture zone'. *Proceedings of the Royal Society, Series B*, **272**: 2525-2533.

Winnard, P., Sidell, B. D. & Vayda, M. E. (2002) Teleost introns are characterized by a high A+T content. *Comparative Biochemistry and Physiology B-Biochemistry and Molecular Biology*, **133**: 155-161.

CHAPTER FOUR

THE PHYLOGENETIC UTILITY OF *TEKTIN*, A NOVEL REGION FOR INFERRING SYSTEMATIC RELATIONSHIPS AMONGST LEPIDOPTERA

Abstract

Rapidly evolving, nuclear coding sequences are highly desirable for phylogenetic studies of closely related species. Here, we investigated an 807 base pair region, homologous to the testis-specific *Tektin* gene from *Bombyx mori*, in 34 nymphalid butterfly taxa in the subfamilies Ithomiinae, Danainae and Heliconiinae. Within Ithomiinae relationships inferred from *Tektin* sequence data were remarkably similar to those in trees based on combined morphological and ecological data. Partitioned Bremer analysis, with mitochondrial *cytochrome oxidase I* and *II*, and nuclear *wingless* and *elongation factor 1-α* sequences, revealed *Tektin* to have the greatest utility for inferring relationships at the genus, tribe and subfamily levels across the studied taxa. We believe *Tektin* will provide a useful source of molecular characters for inference of relationships among other butterflies, and perhaps among other insect taxa.

Introduction

Since the late 1980s, molecular sequence data have played a growing role in phylogenetic analyses. In insects, relatively few loci have contributed to this effort, among them 18S and 28S rDNA, *elongation factor 1-α* (*Ef1α*), *wingless* (*wg*), and the mitochondrial genes *cytochrome oxidase I* and *II* (COI and COII), 12S and 16S. Because gene trees may deviate from the organismal topology due to ancestral polymorphism (Takahata & Nei 1985, Neigel & Avise 1986, Nei 1987, Edwards & Beerli 2000), horizontal gene transfer (Nei 1987, Kidwell 1993, Cummings 1994, Syvanen 1994, Philippe & Douady 2003), and introgressive hybridization (Nei 1987, Doyle 1992, Machado et al. 2002, Mallet 2005), as well as stochastic processes, the sequencing of multiple, unlinked loci is considered highly desirable for obtaining robust phylogenetic inferences that

approximate the organismal phylogeny (Pamilo & Nei 1988, Brower et al. 1996, Edwards & Beerli 2000, Rokas et al. 2003). Yet despite general recognition of these principles, many phylogenetic hypotheses are still based on a single (often mtDNA), or a few separate and putatively independent gene regions.

Mitochondrial sequences are widely used in studies of closely related species, or even subspecies or populations within species, largely because they have a more rapid substitution rate, lower or absent recombination, and smaller effective population size in comparison with nuclear regions (Harrison 1989). However, evidence is mounting that phylogenetic inferences from mitochondrial sequences alone may be confounded by selection (Hudson & Turelli 2003), nuclear copies (Bensasson et al. 2001), and cytoplasmic factors, such as the intracellular bacterium *Wolbachia* (Hurst & Jiggins 2005). Therefore, there exists an urgent need to develop additional nuclear gene sequences that evolve rapidly enough to be used at subspecies, species and higher levels.

We have been investigating phylogenetic relationships among a group of nymphalid butterflies (Ithomiinae) which have long been the subject of research in evolution and biogeography. We sought to develop primers for novel phylogenetically informative nuclear gene regions in these Lepidoptera, amenable to direct sequencing, but which showed rapid evolutionary rates. We deliberately targeted a ~1 kb intronless, easily alignable coding sequence, which could be directly sequenced from genomic DNA using external primers. We repeatedly identified sequences from Ithomiinae inferred to be homologous to a testis-specific *Tektin* from the silk moth, *Bombyx mori* (Genbank accession AB056651) (Ota et al. 2002), while attempting to clone a fragment of *triose-phosphate isomerase* (Beltrán et al. 2002). Initial comparisons with Genbank sequences suggested the potential phylogenetic utility of *Tektin*, so we further investigated the region in 34 species of the subfamilies Ithomiinae, Danainae and Heliconiinae (Nymphalidae), sampled at a range of taxonomic ranks (Table 1). We refer here to Ithomiinae as a subfamily distinct from Danainae, following the most recent checklist for this group by Lamas (2004), although the two are clearly sister taxa and have been considered a single subfamily (Ackery et al. 1999, Brower 2000). Taxa for this study were selected from among those for which comparative data had already been obtained for several other frequently

used Lepidoptera phylogenetic markers: the mitochondrial loci COI and COII, and the nuclear loci *wg* and *Ef1 α*.

Tektins are a family of largely (approx. 70% of their sequence) alpha-helical proteins, which form filaments 2-3 nm dia. (Linck et al. 1985). The filaments are composed of a core heterodimer of Tektins A and B (Pirner & Linck 1994), with Tektin C homodimers either separate, or on the periphery of the Tektin A and B core (Norrander et al. 1996). The filaments are the primary constituent of at least one of 13 protofilaments in the A microtubule wall of doublet microtubules (Linck et al. 1985, Nojima et al. 1995) and are located near to the junction where the A and B microtubules bind (Linck 1990, Nojima et al. 1995). Tektins were first characterised in the purple sea urchin, *Strongylocentrotus purpuratus* (Linck 1982, Linck & Langevin 1982, Linck et al. 1982), which now boasts a well studied Tektin repertoire, having sequence data reported for *Tektin* A1, B1, and C1 (Norrander et al. 1992, Chen et al. 1993, Norrander et al. 1996). *Tektin* GenBank entries are currently available for a number of species: sea urchin, human, mouse, rat, dog, zebra fish, sea squirt, green alga, nematode, fruit fly and silk moth.

The silk moth, *Bombyx mori*, Tektin (GenBank AB056651) had 30% identity to the deduced amino acid (AA) sequence of purple sea urchin Tektin A1, and 28% identities to B1 and C1 over a 338 AA section (AA 159- 496) (Ota et al. 2002). Ota and colleagues (2002) further reported a 38% AA identity of this 338 AA section of the silk moth Tektin to a then hypothetical protein sequence from the fruit fly *Drosophila melanogaster* (AAF44971). This fruit fly sequence remains the closest, non-Lepidoptera sequence to the silk moth Tektin sequence known to date. It is identical to gene product of the more recently submitted *Tektin-A* (NM_078853), suggesting that the only silk moth sequence reported to date is orthologous to *Tektin-A*. The 338 AA silk moth region also has 24% AA identity to the fruit fly Tektin-C (NM_079216) and 26% AA identity to a cloned fruit fly gene product (BT010079). The fruit fly Tektin-A, Tektin-C and cloned product (BT010079) have approximately 25-30% pairwise polypeptide identity over the region corresponding to silk moth aa 159-496, suggesting that the cloned product (BT010079) also belongs to the *Tektin* gene family and may represent the, as yet unassigned, third *D. melanogaster Tektin* gene family member (Tektin-B).

Our major goal was to develop region-specific primers for nuclear genes to extend the currently limited repertoire of gene regions useful for Lepidoptera phylogenetics. Here, we investigate the pattern of molecular evolution of a *Tektin* gene and investigate its utility for inferring relationships amongst Lepidoptera at different taxonomic levels, as compared to expectations based on traditional classification and empirical data from other gene regions.

Materials and methods

DNA extraction

DNA was extracted from one third of a thorax (*Hyposcada* and *Oleria*) or two meso- and metathoracic legs (*Paititia*) using the DNAeasy kit (QIAgen), according to the manufacturer's instructions, with an initial 3 h incubation at 55°C, and a final elution volume of 300 µl. Dried wings were retained as vouchers at University College London. All other samples had previously been isolated from thorax, thorax and head, or thorax, head and abdomen tissue using the SDS phenol/chloroform method (Brower 1994). Dried wings, antennae, legs and available abdomens were retained as vouchers by AVZB at Oregon State University.

Primer development, PCR and sequencing

An individual of *Hyposcada anchiala mendax* (02-1645) (Ithomiinae) was used for initial primer development. Two pairs of primers (Table 2) were designed to amplify overlapping regions. Primers Tek2 and Tek5 were designed from the alignment of silk moth (AB056651) and fruit fly (accession code NM_078853) *Tektin*. Primers Tek3 and Tek4 were designed from a 413 base pair (bp) *Hyposcada* fragment (AY848712) with high resemblance [99 of 121 identities (81%)] to silk moth *Tektin*. The 413 bp putative *Tektin* fragment had been repeatedly amplified in *Hyposcada* samples when attempting to amplify *triose-phosphate isomerase* using TPI-1 and TPI-2 primers (Beltrán et al. 2002).

A PCR was performed in a 50 µl volume, using 4 µl *H. anchiala mendax* (02-1645) template DNA and the following conditions: 1 x PCR buffer (0.1 M Tris HCl, 0.5 M KCl, 0.01 volumes Triton X-100), 0.2 µM dNTPs, 5 µM $MgCl_2$, 0.2 µM each primer, 0.3 U/µl Taq, and an amplification profile of 94°C for 2

Table 1. Specimen and collection locality information, together with the GenBank accession codes for *Tektin* (807 bp), *wg* (417 bp), *Ef1α* (1072 bp) and COI-tRNA-COII mtDNA (1618 bp) sequences. - represents missing data. All nymphalid *Tektin* sequences are first published here, whilst all comparative data were previously generated. The 30 *Hyposcada* and *Oleria*, *wg*, *Ef1α* and mtDNA sequences were obtained from Chapter 6 of this thesis (♠ Whinnett, direct submission 2005). The 4 outgroup sequences, and the other 70 (i.e. non-Oleriini) nymphalid *wg*, *Ef1α* and mtDNA sequences were obtained from GenBank. 55 of these had previously been published, and are indicated as follows: Ω Kamiie et al. 1993; □ Brower 1994; ■ Brower & De Salle 1998, Brower 2000; ○ Beltrán et al. 2002; • Ota et al. 2002; ◊ Brower & Jeansonne 2004; ♣ Brower et al. 2006. Whilst 19 of the 74 sequences obtained from Genbank were direct submissions, and are indicated as follows: ∞ Amanai, direct submission 1999; + Lu et al., direct submission 2001; ♥ Beltrán, direct submission 2005; ♦ Brower, direct submission 2005; and ▲ Zimmermann, direct submission 2005. These different sources sometimes resulted in the different gene regions representing a given species being obtained from more than 1 individual, e.g. *Athyrtis mechanitis salvini* and *Elzunia pavonii*. Individual 02-1645 has 2 *Tektin* GenBank entries, to cover the 2 distinct amplifications: the 412 bp putative *Tektin* region, cloned during initial primer development (AY848712); and the 1142 bp product of direct amplification using the final primers (AY848711).

Taxon	Voucher number	Collection locality	*Tektin*	*wg*	*Ef1α*	mitochondrial
Bombycidae						
Bombyx mori			AB056651 •	D14169 ∞	D13338 Ω	AY048187 +
Nymphalidae: Ithomiinae						
Athyrtis mechanitis salvini Srnka, 1884	RB359	BRAZIL: Rondônia, Cacaulândia	AY848713	DQ071873 ♣	-	DQ069230 ♣
	02-3207	PERU: Cuzco, Palma Real	-	-	DQ085444 ▲	-
Callithomia lenea epidero (Bates, 1862)	RB380	BRAZIL: Rondônia, Cacaulândia	AY848714	-	-	-
Callithomia lenea zelie (Guérin-Méneville, 1844)	PE18-2	PERU: Madre de Dios, Tambopata Preserve	-	DQ071874 ♣	DQ073024 ♣	DQ069232 ♣
Dircenna dero (Hübner, 1823) ssp	E44-3	ECUADOR: Sucumbíos, El Recodo	AY848715	DQ073010 ♣	DQ073025 ♣	DQ069233 ♣
Elzunia pavonii (Butler, 1873)	E28-2	ECUADOR: Loja, San Pedro de Vilcabamba	AY848716	-	-	-
	E12-1	ECUADOR: Loja, Macará	-	AF246562 ■	DQ073026 ♣	DQ069234 ♣
Godyris zavaleta matronalis (Weymer, 1883)	E44-1	ECUADOR: Sucumbíos, El Recodo	AY848717	DQ073011 ♣	DQ073027 ♣	DQ069235 ♣
Greta hermana joiceyi (Kaye, 1918)	E39-46	ECUADOR: Sucumbíos, La Bonita	AY848718	DQ073012 ♣	DQ073028 ♣	DQ069236 ♣
Hyalyris antea amarilla Vitale & Bollino, 2000	E30-4	ECUADOR: Zamora-Chinchipe, Quebrada Chorillos	AY848719	DQ073013 ♣	DQ073029 ♣	DQ069237 ♣
Hyposcada anchiala c. *interrupta* (Tessmann, 1928)	02-2105	PERU: San Martín, km 7.2 Pongo-Barranquita	AY848720	DQ085433 ♠	DQ085445 ♠	DQ078356 ♠
Hyposcada anchiala fallax Staudinger, 1884	02-3519	PERU: Cuzco, Mazuko	AY848721	DQ085434 ♠	DQ085446 ♠	DQ078477 ♠
Hyposcada anchiala mendax Fox, 1941	02-1645	PERU: San Martín, Puente Serranayacu	AY848711/ AY848712	-	-	-
	02-1644	PERU: San Martín, Puente Serranayacu	AY848722	DQ085435 ♠	DQ085447 ♠	DQ078357 ♠
Hyposcada zarepha flexibilis (Haensch, 1909)	02-198	PERU: San Martín, Chumía	AY848723	DQ085436 ♠	DQ085448 ♠	DQ078363 ♠
Ithomia drymo Hübner, 1816	B16-5	BRAZIL: São Paulo, São Paulo	AY848724	DQ073014 ♣	DQ073030 ♣	DQ069238 ♣
Methona sp	RB296	BRAZIL: Rondônia, Cacaulândia	AY848725	DQ073015 ♣	DQ073031 ♣	DQ069239 ♣
Melinaea menophilus Hewitson, 1856	RB239	BRAZIL: Rondônia, Cacaulândia	AY848726	-	-	
	RB288	BRAZIL: Rondônia, Cacaulândia	-	-	-	DQ069240 ♣
	02-1541	PERU: San Martín, Shapaja	-	DQ085437 ▲	-	-
	02-613	PERU: San Martín	-	-	DQ085449 ▲	-

Table 1; continued.

Taxon	Voucher number	Collection locality	*Tektin*	*wg*	*Ef1α*	mitochondrial
Nymphalidae: Ithomiinae						
Napeogenes larilla (Hewitson, 1877) ssp	E39-47	ECUADOR: Sucumbíos, La Bonita	AY848727	DQ073016 ♣	DQ073033 ♣	DQ069241 ♣
Oleria assimilis (Haensch, 1903) ssp nov	02-3609	PERU: Cuzco, Quincemil	AY848728	DQ085438 ♠	DQ085450 ♠	DQ085456 ♠
Oleria gunilla lota (Hewitson, 1872)	Ec467	ECUADOR: Napo, Jatun Sacha	AY848729	DQ085439 ♠	DQ085451 ♠	DQ085457 ♠
Oleria onega janarilla (Hewitson, 1863)	Ec277	ECUADOR: Pastaza, Communidad Shuar Mirador, 70km E of Macas-Puyo	AY848730	DQ085440 ♠	DQ085452 ♠	DQ085458 ♠
	02-515	PERU: San Martín, km 7.2 Pongo-Barranquita	AY848731	DQ085441 ♠	DQ085453 ♠	DQ078390 ♠
Oleria onega (Hewitson, 1852) ssp nov	02-835	PERU: San Martín, Puente Serranayacu	AY848732	DQ085442 ♠	DQ085454 ♠	DQ085459 ♠
Oleria rubescens (Butler & Druce, 1872)	8369	PANAMA: Chiriqui, Quebrada Hornito	AY848733	DQ085443 ♠	DQ085455 ♠	DQ085460 ♠
Pagyris cymothoe sylvella (Hewitson, 1868)	E16-2	ECUADOR: Pichincha, San Antonio	AY848734	-	DQ073035 ♣	DQ157528 ♣
Paititia neglecta Lamas, 1979	02-1244	PERU: San Martín, km 8 Tarapoto-Yurimaguas	AY848735	DQ073017 ♣	DQ073034 ♣	DQ073038 ♣
Pteronymia veia linzera (Herrich-Schäffer, 1864)	E43-16	ECUADOR: Sucumbíos, km 19 La Bonita-Tulcán	AY848736	DQ073018 ♣	DQ073036 ♣	DQ069242 ♣
Tithorea harmonia c. *martina* Fox, 1956	PE19-19	PERU: Huánuco, Tingo María	AY848737	-	-	-
	PE12-3	PERU: Cuzco, Rosalina	-	-	DQ073037 ♦	-
Tithorea harmonia furia (Staudinger, 1884)	V20	VENEZUELA: Monagas, nr Barrancas	-	AF246561 ■	-	DQ157546 ♣
Velamysta pupilla greeneyi Vitale & Bollino, 2003	E43-3	ECUADOR: Sucumbíos, km 19 La Bonita-Tulcán	AY848738	DQ073020 ♣	DQ073021 ♣	DQ071864 ♣
Nymphalidae: Danainae						
Danaus gilippus (Cramer, 1776)	AZ1-3	USA: Arizona, Portal and vicinity	AY848739	DQ175476 ♦	DQ071871 ♦	-
Danaus plexippus (Linnaeus, 1758)	C3-8	COLOMBIA: Meta, Villavicencio, Carretera El Amor	-	-	-	AY569150 ◊
Parantica melusine (Grose-Smith, 1894)	PNG2-3	PAPUA NEW GUINEA: Miamafu	AY848740	-	DQ071870 ♦	DQ175477 ♦
Anetia briarea numidia Hübner [1823]	CU1	CUBA: Santiaga de Cuba	AY848741	AF246579 ■	DQ071869 ♦	DQ071866 ♦
Amauris tartarea Mabille, 1876	GH-058	GHANA: Bobiri Forest Preserve	AY848742	DQ071872 ♦	DQ071868 ♦	DQ071867 ♦

Table 1; continued.

Taxon	Voucher number	Collection locality	*Tektin*	*wg*	*Ef1α*	mitochondrial
Nymphalidae: Heliconiinae						
Philaethria dido (Linneaus, 1763)	P7-4	PANAMA: Colón, Gamboa	AY848743	-	-	U08554 □
	RB283	BRAZIL: Rondônia, Cacaulândia	-	AF014137 ■	-	-
	690	PANAMA: Gamboa, El Renacer	-	-	AY747979 ♥	-
Heliconius sara (Fabricus, 1793) ssp	G42-2	FRENCH GUIANA: Laurent du Maroni, St. Laurent du Maroni	AY848744	-	-	
	P1-7	PANAMA: Colón, Gamboa	-	AF014130 ■	-	-
	308	FRENCH GUIANA: St Maurice on road to Apatou	-	-	AY747924 ♥	-
	STRI-B-850	PANAMA: Canal Zone Pipeline Road	-	-	-	AF413695 ○
Heliconius erato petiverana Doubleday, 1847	P29-2	PANAMA: Panamá, El Llano	AY848745	-	-	-
	CR13	COSTA RICA: Puntarenas, Sirena	-	AF014127 ■	-	-
	2981	PANAMA: Canal Zone, Pipeline Road	-	-	AY748017 ♥	-
	STRI-B-2980	PANAMA: Canal Zone Pipeline Road	-	-	-	AF413684 ○
Heliconius erato (Linneaus, 1758) ssp	P32-7	PANAMA: Darién, Cañazas	AF848746	-	-	-
Heliconius erato hydara Hewitson, 1867	G-16-4	FRENCH GUIANA: Cayenne, Piste Coralie	-	AF014126 ■	-	-
	440	FRENCH GUIANA: Sablance (La Victoire) Route N1 km 19	-	-	AY747987 ♥	-
	STRI-B-442	FRENCH GUIANA: Sablance (La Victorie) Route N1 Km 20	-	-	-	AF413687 ○

min, followed by 32 cycles of (94°C for 60 sec, 50°C for 60 sec, 72°C for 90 sec) and a final 10 min extension at 72°C. PCR products were excised from an agarose gel and purified using the QIAquick gel extraction kit (QIAgen), according to the manufacturer's protocol and sent to a commercial facility for cycle sequencing, precipitation and sequencing. This resulted in a 1142 bp edited product (after loss of data at 3′ end) (AY848711), with a single G↔C difference from the original 413 bp amplification. Final pairwise primers; TektinA, TektinB, Tektin3 and Tektin4, were designed from an alignment of the new 1142 bp *Hyposcada* sequence and the silk moth *Tektin* sequence (Table 2). Using these four primers, the *Tektin* region was amplified and sequenced in 26 Ithomiine species (representing all major Ithomiinae tribes), 4 Danainae, and 4 Heliconiinae species (Table 1). The PCR parameters were as described above, but with the following modifications: 2.5 μM $MgCl_2$, 0.5 μM of forward and reverse primer, a 35 cycle program and a 55°C annealing temperature.

Table 2. Primers designed herein to amplify *Tektin*. Positions are given relative to silk moth *Tektin* (AB056651).

Primer	Sense or Antisense	5′ to 3′ sequence	Nucleotide position
Tek2	AS	TGTCRCTCCAATCRWATTC	930-948
Tek3	S	CAAGACGTACAAGCTAGCAAGA	198-219
Tek4	S	ACTGGAGAATGGCGAAAGAAC	459-479
Tek5	AS	GGCCWMRTCKCKGCAGTT	1401-1418
TektinA	S	ACCAGTGGRGAYATYCTWGG	330-349
TektinB	S	CAGGMCAYATMGAYTGGA	376-393
Tektin3	AS	CGCAGTTTYTGATRCTYT	1084-1101
Tektin4	AS	TCATRTCTTGASWGCCTTTG	1157-1176

Data analysis

Tektin sequences were edited using SeqEd v1.0.3 software (Applied Biosystems, Inc), and peptide sequences generated in MacClade 4.0 (Maddison & Maddison 1997). Sequences were deposited in Genbank (Accession numbers listed in Table 1). All comparative (*Ef1 α*, *wg* and mitochondrial), and *Bombyx mori* (outgroup) sequences were obtained from previous publications and/or other research projects, as detailed in Table 1. PAUP version 4.0b 10 (Swofford 2000) was used to calculate: numbers of variable and parsimony informative sites, uncorrected pairwise divergences, nucleotide composition, transition and

transversion statistics, and to perform a chi-square test of base frequency homogeneity. Data partition homogeneity tests (Farris et al. 1994) were also implemented in PAUP between all pairwise combinations of gene partitions in all individuals for which complete or nearly-full length data were available for both gene partitions. Phylogenetic analyses were performed using MrBayes 3.0 (Huelsenbeck & Ronquist 2001), with nst=6 and a gamma rates heterogeneity model, with four simultaneous chains run for 1,000,000 generations, sampling a tree every 100 generations. A consensus tree with branch support in the form of posterior probabilities was derived from the final 9,000 trees (representing the final 900,000 generations), after confirmation that likelihood values had stabilized. We also conducted a heuristic search with TBR branch swapping in PAUP to find the most parsimonious trees, and summarized these with a majority rule consensus tree. Parsimony node support was assessed by Bremer support and bootstrapping with 1,000 replicates. To dissect the contributions of the individual genes in recovering relationships at different taxonomic levels, partitioned Bremer support values (Baker & DeSalle 1997) were calculated for the total combined data, using TreeRot (Sorenson 1999). Finally, a phylogenetic hypothesis for the translated sequences with 1,000 bootstrap replicates was estimated using neighbour joining in PAUP.

Results

Of the pairwise final *Tektin* primer combinations tested, primers TektinA and Tektin3 consistently generated the strongest PCR products in all specimens except for Heliconiinae, for which primers TektinB and Tektin4 were optimal. Sequence data generated using TektinA with Tektin3 and TektinB with Tektin4 were edited to give 729 bp, and 752-763 bp intronless regions respectively (removal of ambiguous chromatogram peaks resulted in some loss of data at the 3′ end). Within the 807 bp final alignment, 685 bp of overlap was amplified using both primer pairs, of which 387 nucleotides were variable among all specimens (317 of which were parsimony informative (PI)), 364 variable (310 PI) among all Nymphalidae, 328 variable (257 PI) among Danainae and Ithomiinae, 278 variable (189 PI) within Ithomiinae, and 84 variable (67 PI) among *Oleria* and *Hyposcada* (Ithomiinae: Oleriini). These variable sites are

approximately randomly distributed along the entire length of the amplified region.

Wingless sequences produced a 417 bp final alignment. Partially overlapping mitochondrial and *Ef1 α* regions had been amplified, due to the different primer choices of collaborating laboratories. These sequences were edited to give final alignments of the 1618 bp (mitochondrial) and 1072 bp (*Ef1 α*) which were complete for most taxa. The mitochondrial region has the greatest number of variable sites, for example, 577 (436 PI) within Nymphalidae and 497 (347 PI) among Ithomiinae. However, when accounting for length differences, *Tektin* is more variable, having 45.1% and 34.4% variable sites within Nymphalidae and Ithomiinae respectively, compared to the corresponding values of 35.7% and 30.7% for the mitochondrial data. Both absolutely and relative to length, *Tektin* has higher variability than the other nuclear regions. The *Ef1 α* fragment has 292 (27.2%) variable sites (235 PI) among the Nymphalidae, and 238 (22.2%) (158 PI) among the Ithomiinae. And the shortest region, *wg*, has 163 (39.1%) (130 PI) and 121 (29.0%) (83 PI) variable sites within the Nymphalidae and Ithomiinae respectively.

The amplified *Tektin* region is AT rich (Fig. 1). The sum of the mean contributions of A and T nucleotides to the 807 bp alignment is 59.9%, with first, second and third codons having respective, increasing AT biases of 55.3%, 61.8% and 62.5%. The overall AT bias is more pronounced than in *Ef1 α* (49.5%) and *wg* (46.2%), but considerably less than that in mitochondrial DNA (75.1%) in corresponding taxa. Heterogeneity in the nucleotide composition of synonymous sites was observed. For example, 21A, 9C, 8G and 31T nucleotides comprise the 69 four-fold degenerate sites along the *Tektin* region in *Tithorea*. This distribution represents a clear departure from an equal nucleotide composition (1A: 1C: 1G: 1T) (chi square test, $P < 0.001$), as well as a strong departure from the overall nucleotide composition of the entire *Tektin* region in *Tithorea* (264A: 117C: 166G: 179 T) (chi square test, $P < 0.001$). Chi square tests for base frequency homogeneity, for the total *Tektin* dataset, first, second and third codon positions, revealed no significant differences across taxa ($P > 0.95$).

To assess the extent of saturation, the proportion of sites with transitional and transversional pairwise changes were plotted against uncorrected pairwise divergences (Fig. 2). The near linear trends of all transversions, and first and second codon position transitions imply that saturation has not been reached for any of the pairwise comparisons, in any gene. The third position transition trend lines begin to plateau as the more phylogenetically distant taxa are considered, thus deeper evolutionary hypotheses from each dataset might be confounded by multiple substitutions. This plateau effect was strongest for the mitochondrial dataset.

Figure. 1. Average nucleotide composition of the *Tektin* fragment (corresponding to the 807 bp region between primers TektinA and Tektin4), in the 35 taxa (1 bombycid and 34 nymphalid) studied here. Error bars depict the minimum to maximum contribution of each nucleotide, as observed in individual specimens.

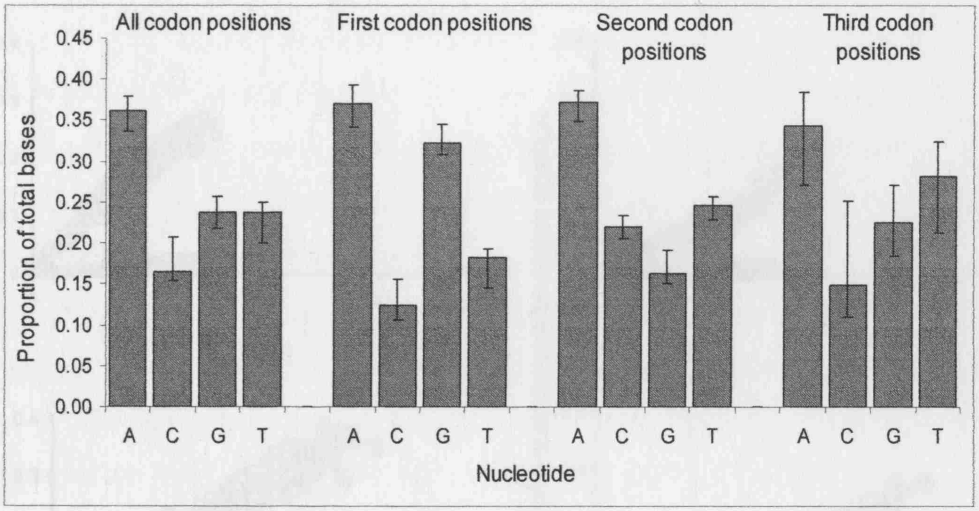
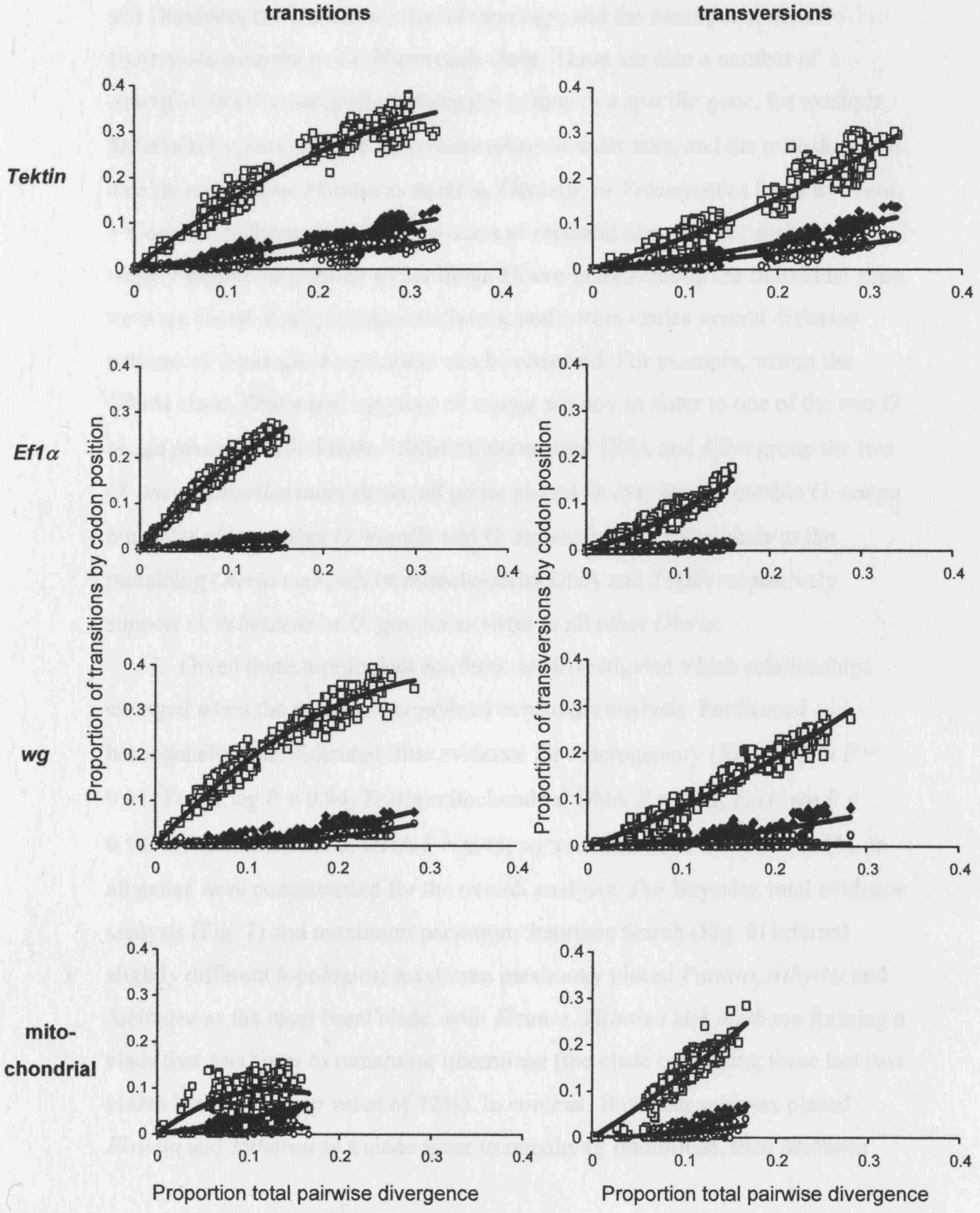


Table 3. The number of Bayesian inferred branches with strong Bayesian, Parsimony, or Bayesian and Parsimony support. Branches are regarded as strongly supported if they have a Bayesian posterior probability greater than 0.70, or bootstrap value greater than 70% in the parsimony analysis. Mean Bayesian posterior probability and bootstrap values are given in parentheses.

	Bayesian	Parsimony	Bayesian + Parsimony
Tektin	4 (0.90)	-	22 (0.99, 92%)
wg	4 (0.83)	2 (75%)	6 (0.93, 92%)
Ef1α	8 (0.87)	-	20 (1.00, 92%)
mitochondrial	11 (0.93)	-	6 (1.00, 97%)

Figure. 2. Proportion of sites with transitional or transversional changes, against uncorrected pairwise divergence for: the 807 bp *Tektin* fragment corresponding to the region between primers TektinA and Tektin4; as well as the 1072 bp *Ef1α*; 417 bp *wg*; and 1618 bp mtDNA regions studied here. ♦ represents the first, O represents the second, and □ represents the third codon positions. Best-fit logarithmic trendlines were added using Excel. Near-linear trendlines imply that saturation has not been reached, whereas a plateau (as observed for mtDNA third position transitions) suggests that data from these sites might be confounded by multiple hits. Note that the *Ef1α* first and second codon position transversion markers are hard to distinguish as they share a similar trajectory.



The different genes gave strongly heterogeneous support for phylogenetic hypotheses: six *wg* and mitochondrial, 20 *Efl α* and 22 *Tektin* branches were supported by both a Bayesian posterior probability greater than 0.70 and a bootstrap value greater than 70% in the parsimony analysis (Figs. 3-6.; and Table 3). There are cases of agreement between the phylogenetic hypotheses inferred by Bayesian methods across all genes, including monophyly of both Heliconiinae and Danainae, the Danainae internal topology, and the basal position of *Hyposcada zarepha* in the *Hyposcada* clade. There are also a number of examples where a particular topology is unique to a specific gene, for example, *Efl α* fails to place the two *Heliconius erato* as sister taxa, and the mitochondrial data do not recover *Elzunia* as sister to *Tithorea*, or *Velamysta* as basal to *Greta* + *Godyris*. Additionally, there are cases of repeated discordance, such as the weakly supported position of *Methona*. Discordances among the individual gene trees are found at all phylogenetic levels, and within clades several different patterns of topological agreement can be observed. For example, within the *Oleria* clade, *Tektin* and *wg* place *O. onega* ssp nov as sister to one of the two *O. onega janarilla* individuals, whilst mitochondrial DNA and *Efl α* group the two *O. onega janarilla* individuals; all genes placed *O. assimilis* as outside *O. onega,* but *Efl α* and *wg* place *O. gunilla* and *O. rubescens* as a sister clade to the remaining *Oleria* taxa, whilst mitochondrial DNA and *Tektin* respectively support *O. rubescens* or *O. gunilla* as sister to all other *Oleria.*

Given these topological conflicts, we investigated which relationships emerged when the data were combined in a single analysis. Partitioned homogeneity tests indicated little evidence for heterogeneity (*Efl α*:*Tektin* P = 0.91, *Tektin*:*wg* P = 0.94, *Tektin*:mitochondrial DNA P = 0.13, *Efl α*:*wg* P = 0.98, *Efl α*:mitochondrial DNA P = 0.45, *wg*:mitochondrial DNA P = 0.43), so all genes were concatenated for the overall analysis. The Bayesian total evidence analysis (Fig. 7) and maximum parsimony heuristic search (Fig. 8) inferred slightly different topologies; maximum parsimony placed *Paititia*, *Athyrtis* and *Melinaea* as the most basal clade, with *Elzunia*, *Tithorea* and *Methona* forming a clade that was sister to remaining Ithomiinae (the clade containing these last two clades had a bootstrap value of 72%). In contrast, Bayesian analyses placed *Elzunia* and *Tithorea* as a clade sister to remaining Ithomiinae, then *Methona*

(posterior probability 0.56), then the *Athyrtis*, *Paititia*, *Melinaea* clade as sister to remaining Ithomiinae (posterior probability 0.56).

Overall, the total evidence Bayesian topology was most similar to the individual topologies inferred using *Tektin* and *Ef1 α*. The *Tektin* and *Ef1 α* topologies both differed from the total evidence topology in arrangements of the *Paititia*, *Athyrtis* and *Melinaea* clade, and placements of *Oleria rubescens* and *Oleria gunilla*. In addition, the *Tektin* topology differed from the total evidence topology at the intraspecific level (i.e. *Hyposcada anchiala* and *Oleria onega*), whilst *Ef1 α* differed in the arrangement of the *Velamysta*, *Greta*, *Godyris* clade, position of *Methona*, and relationships between *Heliconius* species.

Mitochondrial DNA provides the strongest support for relationships between closely related taxa, as measured by Partitioned Bremer support values within the *Hyposcada* and *Oleria* clades. The Partitioned Bremer support analysis also revealed that *Tektin* most strongly supported the deeper nodes, indicating it contributed significantly to resolving positions of the most divergent taxa in the combined analysis.

Topologies inferred from Tektin protein (Fig. 9) and *Tektin* nucleotide data are in broad agreement, for example, by recovering the monophyly of the three nymphalid subfamilies, the basal positions of *Elzunia*, *Tithorea* and *Methona* within the Ithomiinae clade, and the derived position of the monophyletic Oleriini. *Velamysta* assumes the most discordant position, despite having lowest protein character differences to the taxa with which it clusters based on nucleotide data (14 with *Greta* and 12 with *Godyris*, compared to, for example, 21 with *Napeogenes*).

Figure. 3. Phylogenetic hypothesis based on 729 bp of *Tektin* nucleotide data, corresponding to the region amplified using TektinA and Tektin3 primers. Note that as a result of using TektinB and Tektin4 PCR primers, the first 46 bp of this region were not amplified in the specimens of Heliconiinae, and have therefore been coded as missing data. Tree topology is consensus of the last 9,000 trees inferred using Bayesian methods. Bayesian probabilities above 0.70 (above branch) and parsimony bootstrap support values higher than 70% (below branch) are given. This topology is in broad agreement to traditional classification at a number of levels (Keith R Willmott, personal communication). For example, monophyly of the 3 nymphalid subfamilies, and clustering of the Oleriini, are recovered.

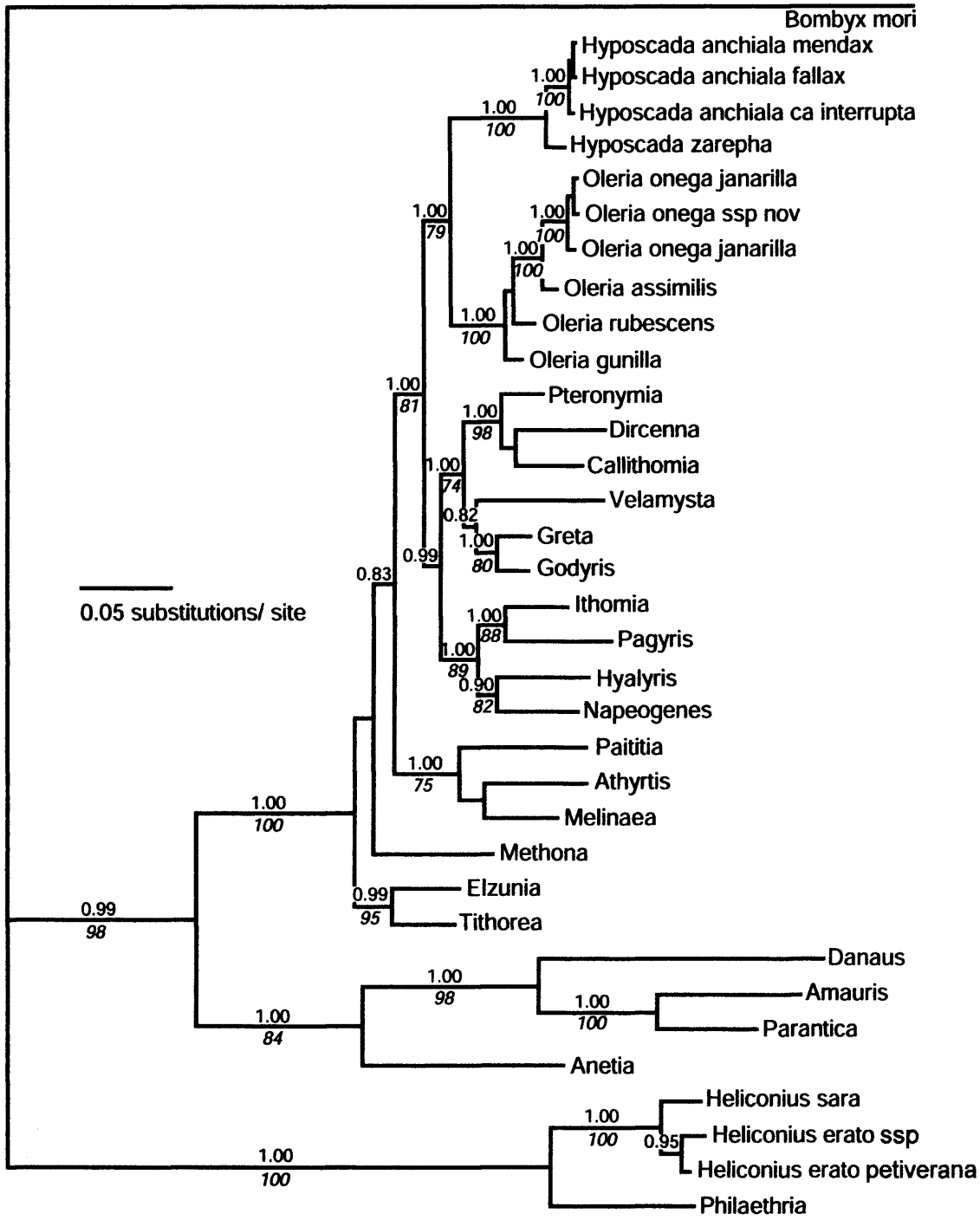


Figure. 4. Phylogenetic hypothesis based on 417 bp *wg* nucleotide data, for representatives of all taxa except *Parantica melusine* (due to missing data). *Tree* topology is consensus of the last 9,000 trees inferred using Bayesian methods. Bayesian probabilities above 0.70 (above branch) and parsimony bootstrap support values higher than 70% (below branch) are given. Note that the paraphyly of the Ithomiine, with respect to the Heliconiinae, as shown here, is in contrast to the subfamily monophyly recovered using morphological characters, as well as *Tektin*, *Ef1α* and mtDNA gene regions.

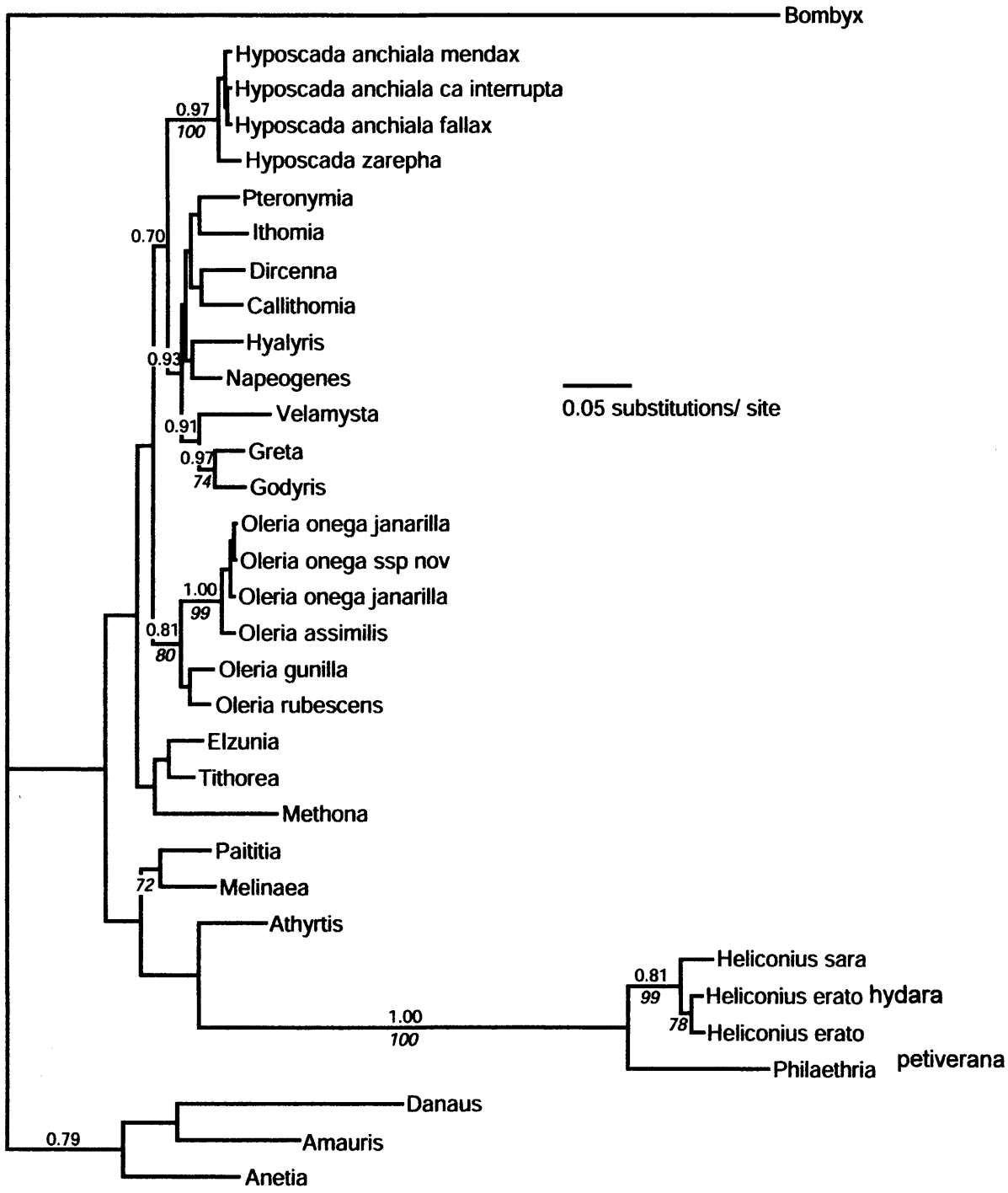


Figure. 5. Phylogenetic hypothesis based on 1072 bp *Ef1α* nucleotide data. Tree topology is consensus of the last 9,000 trees inferred using Bayesian methods. Bayesian probabilities above 0.70 (above branch) and parsimony bootstrap support values higher than 70% (below branch) are given. Note that *Ef1α* does not recover the two *Heliconius erato* subspecies as sister taxa. This is in contrast to the *H. erato* sister relationship proposed using morphological characters, as well as *Tektin*, *wg* and mtDNA gene regions.

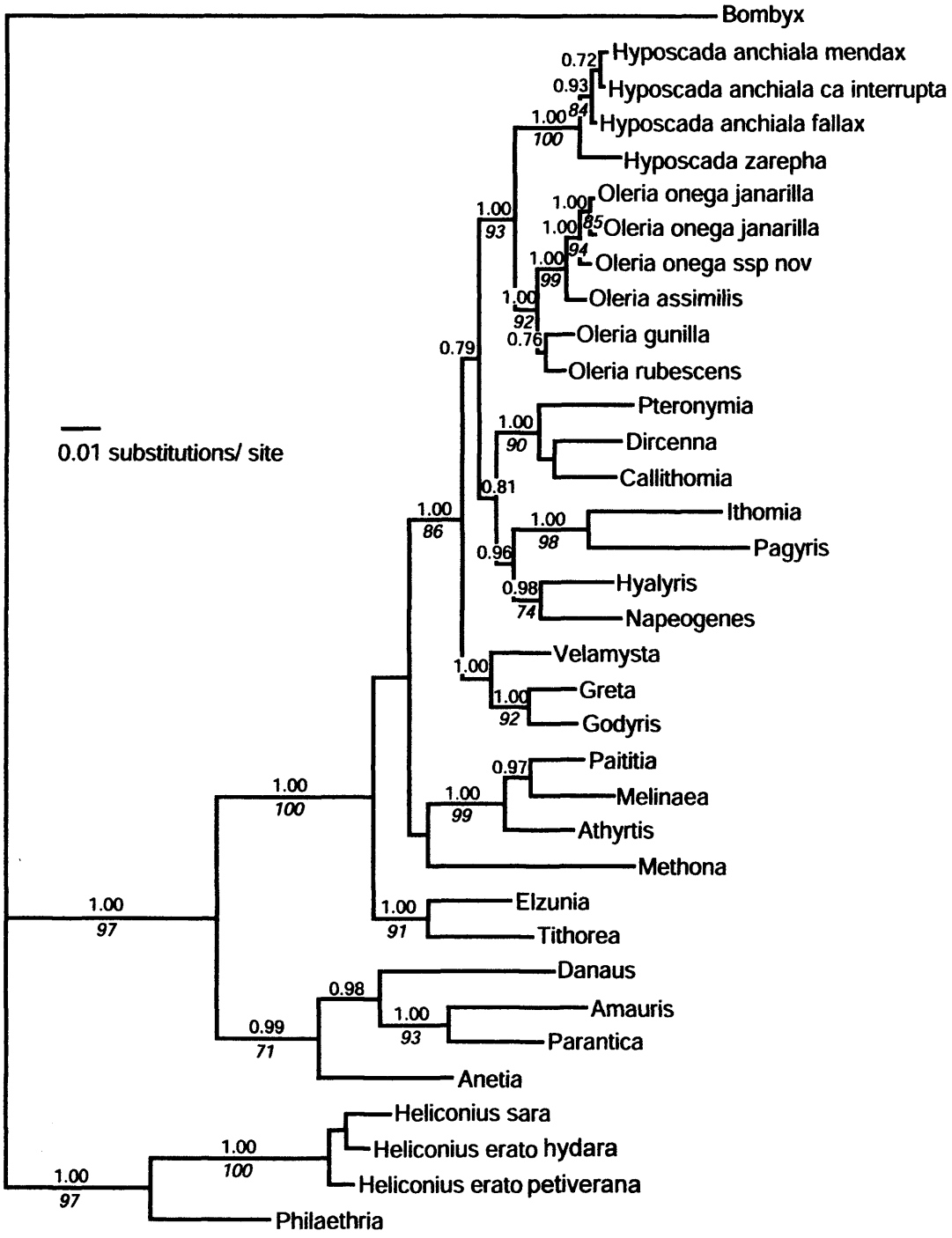


Figure. 6. Phylogenetic hypothesis based on 1618 bp mtDNA data (COI-tRNA-COII). Tree topology is consensus of the last 9,000 trees inferred using Bayesian methods. Bayesian probabilities above 0.70 (above branch) and parsimony bootstrap support values higher than 70% (below branch) are given, and show that the mtDNA data confer strong support to relationships between closely related taxa.

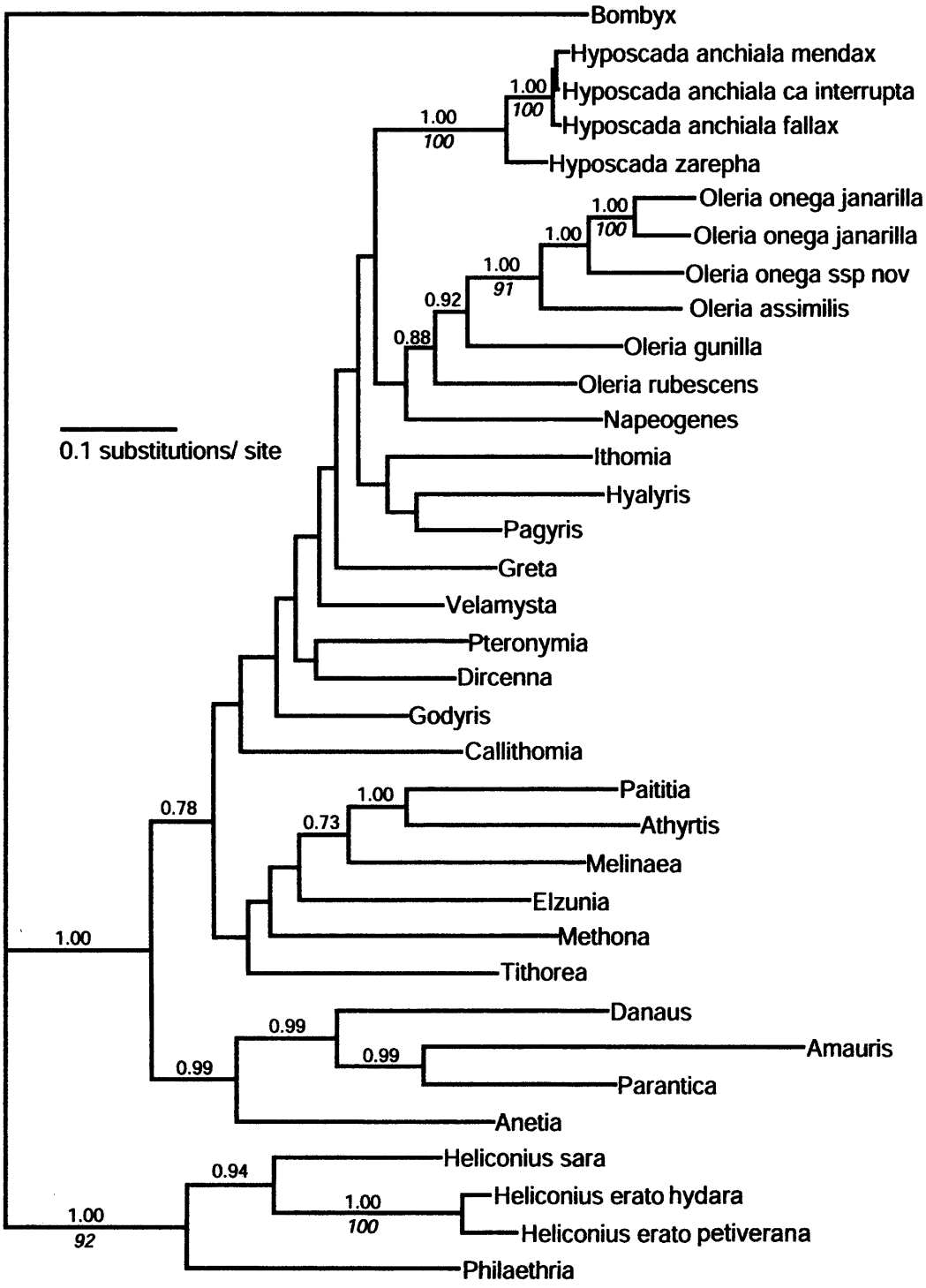


Figure. 7. Total evidence phylogenetic hypothesis. Each taxon is represented by data for *Tektin, wg, Ef1α* and mtDNA regions, except for *Parantica melusine*, which is missing data for the *wg* region. Note that in a number of cases, terminal taxa are represented by gene regions from more than one individual (due to the inclusion of sequences generated by multiple authors, as detailed in Table 1). Tree topology is consensus of last 9,000 trees inferred using Bayesian methods. Bayesian probabilities above 0.70 (above branch) and parsimony bootstrap support values higher than 70% (below branch) are given. This total evidence hypothesis represents the closest, out of those presented in Figs. 3- 8 of this chapter, to that expected on theoretical grounds (DeSalle and Brower 1997).

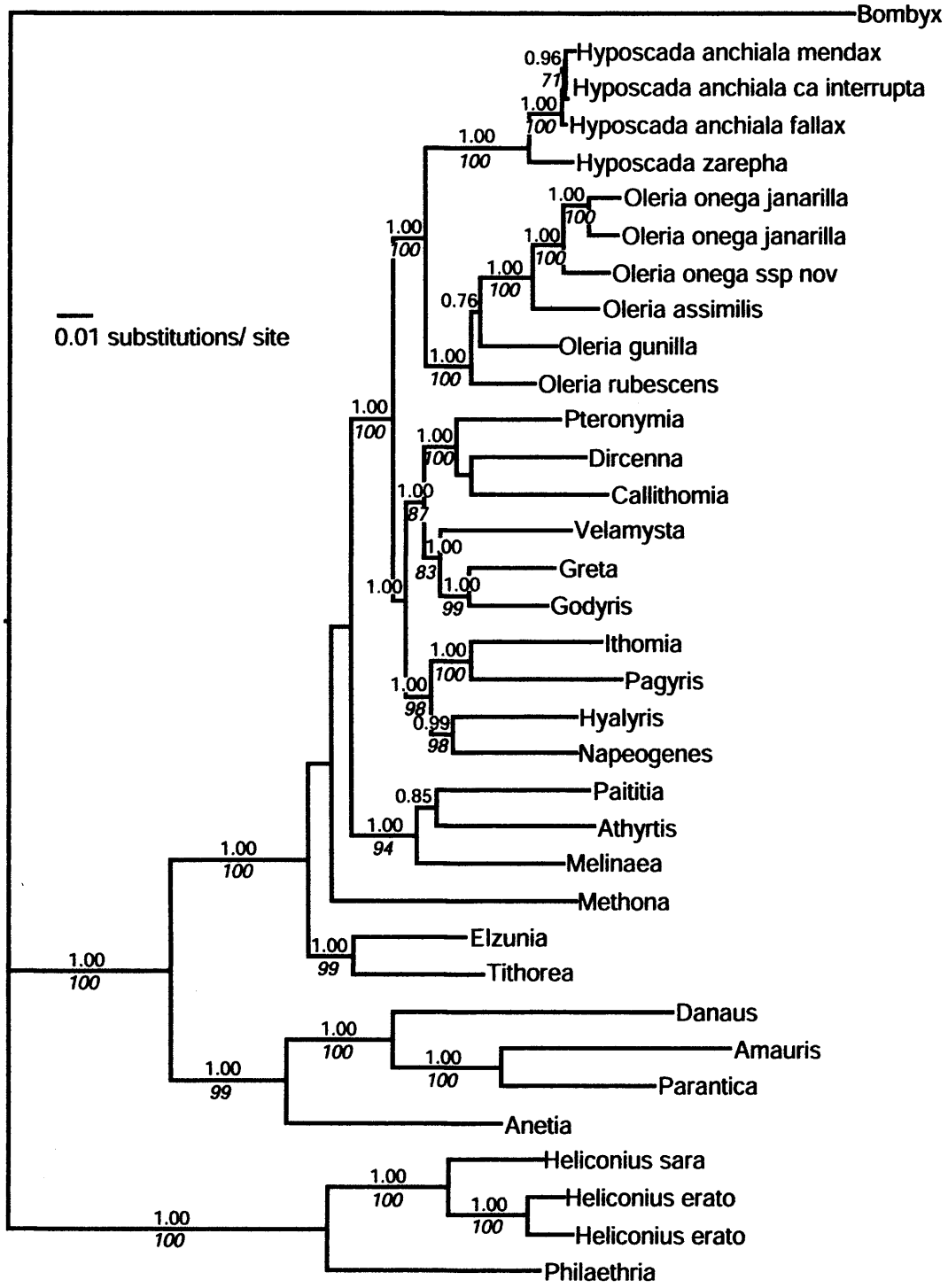


Figure. 8. Total evidence phylogenetic hypothesis. Each taxon is represented by data for *Tektin*, *wg*, *Ef1α* and mtDNA regions, except for *Parantica melusine*, which is missing data for the *wg* region. Note that in a number of cases, terminal taxa are represented by gene regions from more than one individual (due to the inclusion of sequences generated by multiple authors, as detailed in Table 1). Tree topology was inferred with maximum parsimony. The hypothesis presented here differs slightly from the total evidence hypothesis inferred using Bayesian methods (see Fig. 7), for example, in the arrangement of the basal ithomiines. Partitioned Bremer support values are given above branches (*Tektin*/ *Ef1α*/ *wg*/ mitochondrial), and show that the mitochondrial region provides strongest support for relationships between closely related taxa, whilst the *Tektin* region most strongly supports the deeper nodes.

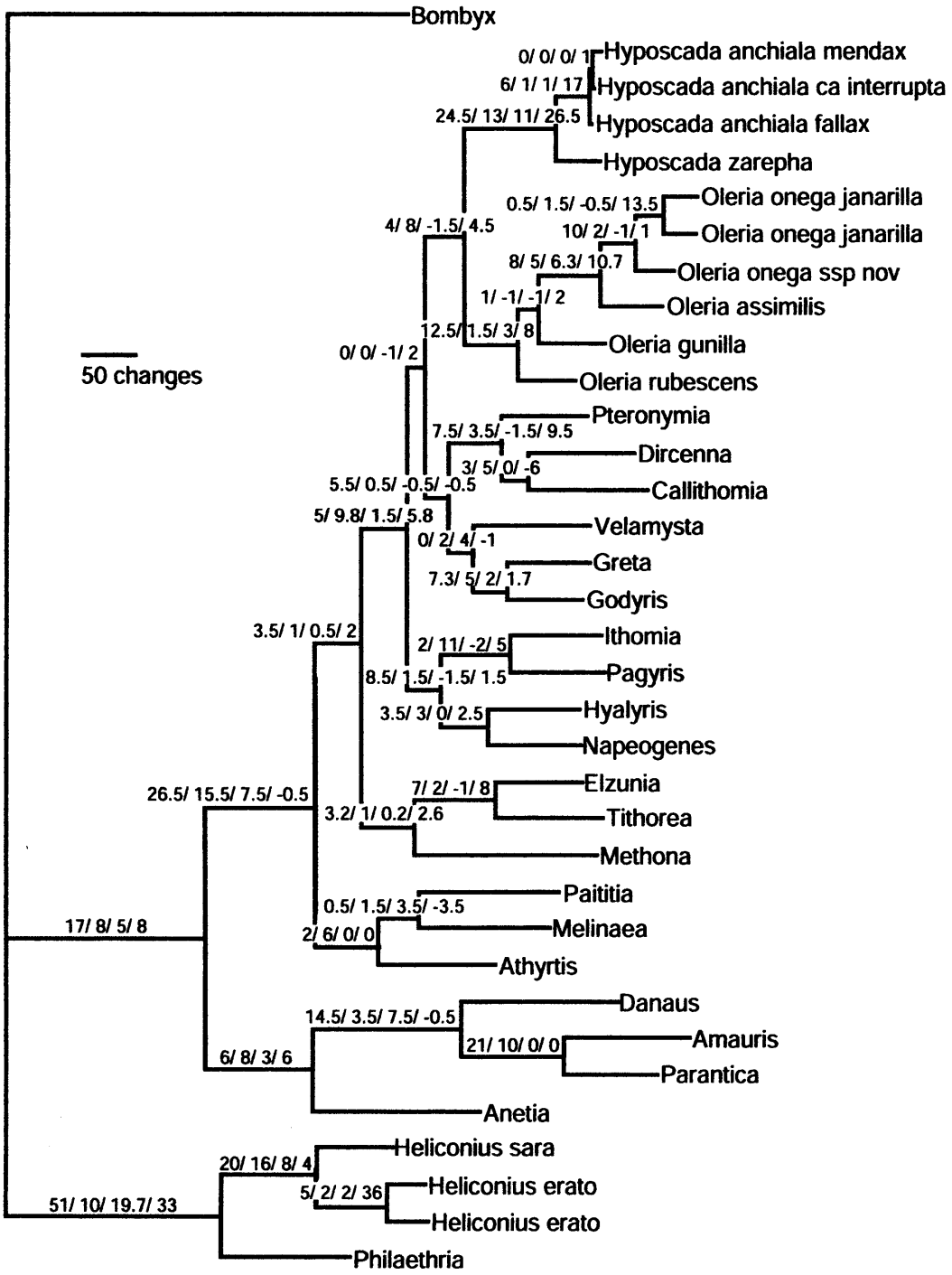
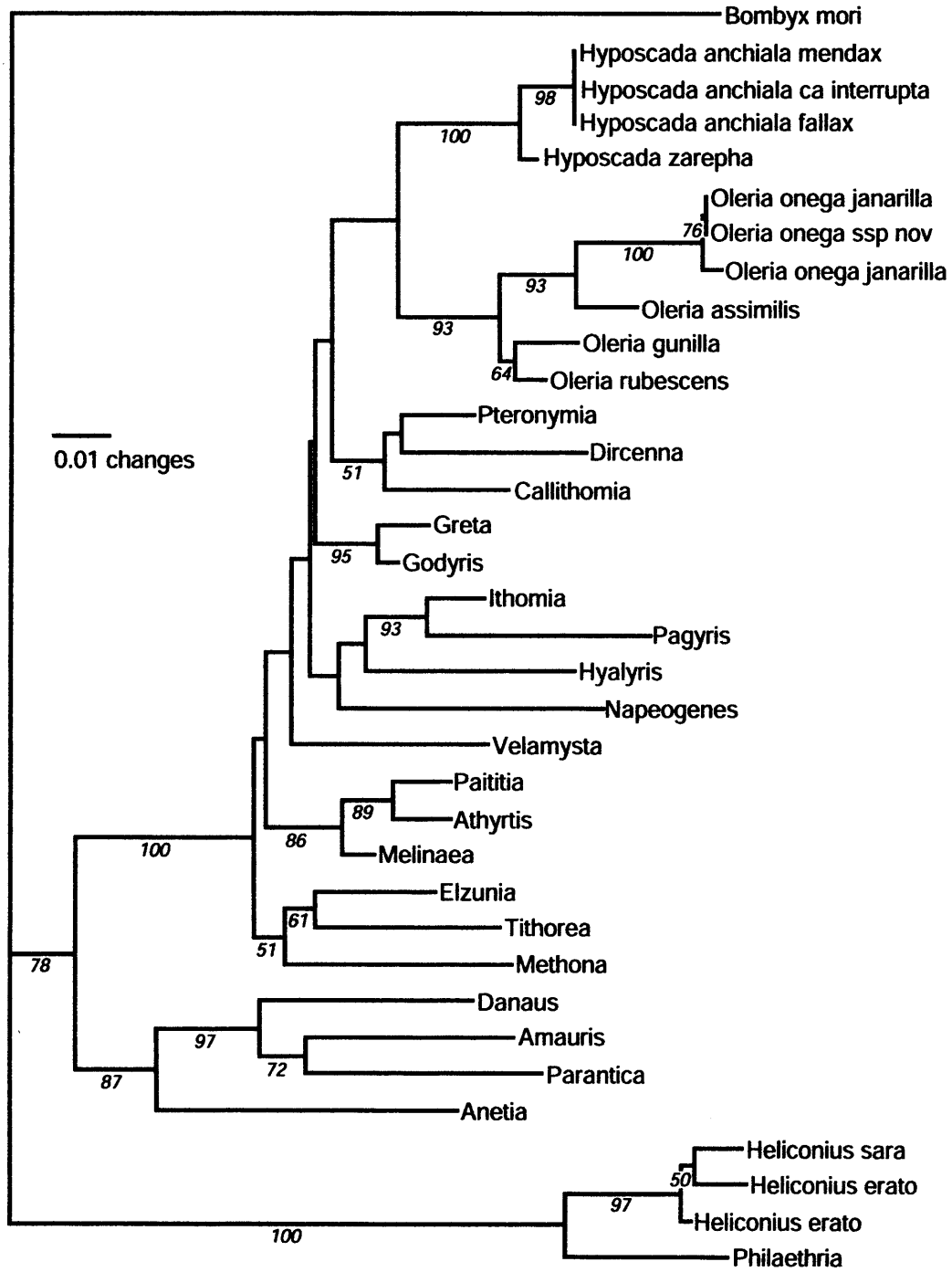


Figure. 9. Phylogenetic hypothesis based on Tektin amino acid data, inferred with neighbour joining. The topology presented here is in broad agreement with the hypothesis based on *Tektin* nucleotide data (see Fig. 3). Neighbour joining bootstrap support values are given below branches.



Discussion

Tektin proved easy to amplify from genomic DNA in all specimens tested, was amenable to direct sequencing, was easily aligned and exhibited a level of variation particularly appropriate for resolving phylogenetic relationships at the genus, tribe and subfamily levels.

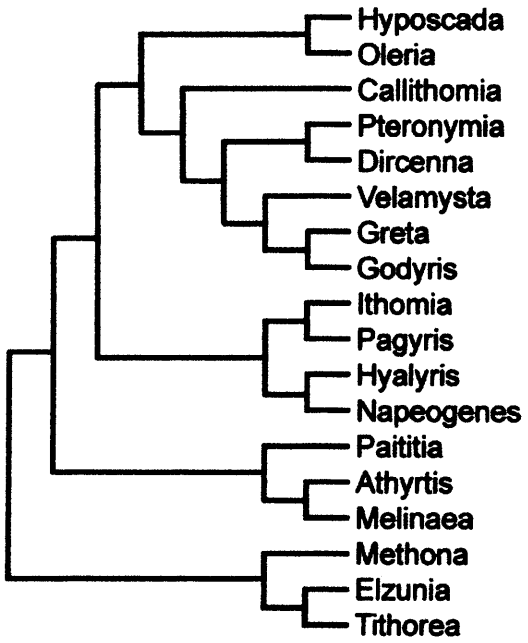
Tektin products belong to a protein family comprising of at least 3 forms (A, B, C), thus the potential for amplifying paralogs was of concern. However, the high nucleotide differentiation of *Tektin* paralogs makes them easily distinguishable, for example, uncorrected pairwise distances of A:B 52%, A:C 48% and B:C 52% were recovered for a 951 bp region of purple sea urchin *Tektin* which aligns without insertions and deletions (M97188, L21838 and U38523). As the pairwise distance values recovered in this study are all significantly lower than the distances observed between *Tektin* paralogs in other species, for example, within ithomiines (0 to 15.2%, mean 10.3%) and even between the silk moth outgroup and all other individuals (26.5 to 32.1%, mean 29.1%), it is highly likely that all our sequences are actually orthologous. In addition, phylogenetic analysis of Genbank *Tektin* sequences in human, mouse, sea urchin, fruit fly, silk moth and our ithomiine sequences (data not shown) shows that *Tektin* sequences cluster together by paralog rather than by the species to which they belong. This clearly indicates that individual *Tektin* paralogs really are distinct enough to behave as single copy genes. Easily distinguishable paralogs have previously proven effective for phylogenetic analyses, including *Efl α* and *wg* (Cho et al. 1995, Brower and De Salle 1998).

The mitochondrial and *Tektin* regions had the highest number of variable sites. However, when accounting for length differences, *Tektin* had the highest per-site variability. Despite its faster evolutionary rate, and the resultant concern that deep level comparisons using highly variable regions are often more confounded by the effects of saturation, *Tektin* proved to be more informative than mitochondrial DNA for the more distant comparisons. This is probably due to the biased nucleotide composition of the mitochondrial region (75.1%, compared to 59.9% in *Tektin*) hindering its deeper resolving power.

That the *Tektin* topology was in broad agreement to traditional classification and a recent cladogram inferred from morphology (Fig. 10, Keith Willmott and André Freitas, personal communication) strongly supports the

utility of *Tektin* as a tool for successfully inferring relationships. The *Tektin* gene tree was most similar to that obtained from the *Ef1 α* dataset, despite the differing placements of *Velamysta*, *Greta* and *Godyris*. Both *Tektin* and *Ef1 α* topologies are well resolved, with most branches well supported. Mitochondrial DNA best differentiated closely related taxa within the Oleriini, but it failed to provide congruent support for intermediate depth internal branches and its placement of *Napeogenes* within Oleriini was clearly erroneous; both effects are probably due to the confounding effects of saturation, a by-product of the limited taxonomic sampling employed in this analysis. The topology generated from the *wg* data has more internal structure than the mitochondrial tree, but is also quite poorly resolved, in part because of short sequence length. These findings are based on our taxa of choice and may not have done justice to each gene, which have previously proven to be useful (Simon et al. 1994, Cho et al. 1995, Brower & De Salle 1998, Caterino et al. 2000). Our purpose was not to advocate or discourage general use of these "paradigm" genes, but to provide comparative evidence to enable future studies to select the combination of genes most appropriate for their level of phylogeny estimate.

Figure. 10. Most parsimonious tree inferred from combined morphological and ecological data, after successive approximations character weighting (Keith Willmott and André Freitas, personal communication).



After combining *Tektin* with COI and COII, *wg* and *Ef1α* regions, partitioned Bremer analyses demonstrated *Tektin* to be particularly effective at recovering the genus, tribe and subfamily relationships. The total evidence tree represents our preferred phylogenetic hypothesis, as is to be expected on theoretical grounds (DeSalle & Brower 1997), and also shows good agreement with the morphological tree (Fig. 10). Given the close correspondence between trees derived from *Tektin* alone and both combined evidence molecular and morphological trees, *Tektin* is arguably the most accurate single marker for inferring relationships among all taxa studied. We recommend future studies employ *Tektin* in conjunction with other genomic regions, in order to obtain robust and reliable phylogenetic hypotheses that best imply the organismal phylogeny. We believe *Tektin* to be a very useful addition to the molecular phylogenetic arsenal for studies of Lepidoptera. Our findings strongly endorse *Tektin* as a new candidate gene for phylogenetic and perhaps phylogeographical studies of other insect groups.

References

Ackery, P. R., De Jong, R. & Vane-Wright, R. I. (1999) The butterflies: Hedyloidea, Hesperioidea and Papilionoidea. Pp. 263-300. In N. P. Kristensen (ed.), Lepidoptera, moths and butterflies. 1. Evolution, systematics and biogeography. *Handbook of Zoology*, 4 (35), Lepidoptera. de Gruyter, Berlin.

Baker, R. H. & DeSalle, R. (1997) Multiple sources of character information and the phylogeny of Hawaiian Drosophilids. *Syst. Biol.*, **46**: 654-673.

Beltrán, M. S., Jiggins, C. D., Bull, V., Linares, M., Mallet, J., McMillan, W. O. & Bermingham, E. (2002) Phylogenetic discordance at the species boundary: comparative gene genealogies among rapidly radiating *Heliconius* butterflies. *Mol. Biol. Evol.*, **19**: 2176-2190.

Bensasson, D., Zhang, D.-X., Hartl, D. L. & Hewitt, G. M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol. Evol.*, **16**: 314-321.

Brower, A. V. Z. (1994) Phylogeny of *Heliconius* butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae). *Mol. Phylogenet. Evol.*, **3**: 159-174.

Brower, A. V. Z. (2000) Phylogenetic relationships among the Nymphalidae (Lepidoptera) inferred from partial sequences of the *wingless* gene. *Proc. R. Soc. Lond. B Biol. Sc.*, **267**: 1201-1211.

Brower, A. V. Z. & De Salle, R. (1998) Patterns of mitochondrial vs. nuclear DNA sequence divergence in nymphalid butterflies: the utility of *wingless* as a source of characters for phylogenetic inference. *Insect Mol. Biol.*, **7**: 73-82.

Brower, A. V. Z., DeSalle, R. & Vogler, A. (1996) Gene trees, species trees, and systematics. A cladistic perspective. *Annu. Rev. Ecol. Syst.*, **27**: 423-450.

Brower, A. V. Z. & Jeansonne, M. M. (2004) Geographical populations and 'subspecies' of New World monarch butterflies (Nymphalidae) share a recent origin and are not phylogenetically distinct. *Ann. Entomol. Soc. Am.*, **97**: 519-523.

Brower, A. V. Z., Freitas, A. V. L., Lee, M.-M., Silva Brandão, K. L., Whinnett, A. & Willmott K. R. (2006) Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions. *Systematic Entomology*, **31**: 288-301.

Caterino, M. S., Cho, S. & Sperling, F. A. H. (2000) The current state of insect molecular systematics: a thriving Tower of Babel. *Annu. Rev. Entomol.*, **45**: 1-54.

Chen, R., Perrone, C. A., Amos, L. A. & Linck, R. W. (1993) Tektin B1 from ciliary microtubules: primary structure as deduced from the cDNA sequence and comparison with Tektin A1. *J. Cell Sci.*, **106**: 909-918.

Cho, S., Mitchell, A., Regier, J. C., Mitter, C., Poole, R. W., Friedlander, T. P. & Zhao, S. (1995) A highly conserved nuclear gene for low-level phylogenetics: *elongation factor 1-alpha* recovers morphology-based tree for heliothine moths. *Mol. Biol. Evol.*, **12**: 650-656.

Cummings, M. P. (1994) Transmission patterns of eukaryotic transposable elements: arguments for and against horizontal transfer. *Trends Ecol. Evol.*, **9**: 141-145.

DeSalle, R. & Brower, A. V. Z. (1997) Process partitions, congruence and the independence of characters: inferring relationships among closely-related Hawaiian *Drosophila* from multiple gene regions. *Syst. Biol.*, **46**: 751-764.

Doyle, J. (1992) Gene trees and species trees: molecular systematics as one-character taxonomy. *Syst. Bot.*, **17**: 144-163.

Edwards, S. V. & Beerli, P. (2000) Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**: 1839-1854.

Farris, J. S., Källersjö, M., Kluge, A. G. & Bult, C. (1994) Testing significance of incongruence. *Cladistics*, **10**: 315-319.

Harrison, R. G. (1989) Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends Ecol. Evol.*, **4**: 6-11.

Hudson, R. R. & Turelli, M. (2003) Stochasticity overrules the "three-times rule": genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, **57**: 182-190.

Huelsenbeck, J. P. & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**: 754-755.

Hurst, G. D. D. & Jiggins, F. M. (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic, and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **272**: 1524-1534.

Kamiie, K., Taira, H., Ooura, H., Kakuta, A., Matsumoto, S., Ejiri, S. & Katsumata, T. (1993) Nucleotide sequence of the cDNA encoding silk gland elongation factor 1 alpha. *Nucleic Acids Res.*, **21**: 742.

Kidwell, M. G. (1993) Lateral transfer in natural populations of eukaryotes. *Annu. Rev. Genet.*, **27**: 235-256.

Lamas, G. (2004) Nymphalidae. Ithomiinae, pp. 172-191. In G. Lamas (ed.), Checklist: Part 4A. Hesperioidea - Papilionoidea. In J. B. Heppner, (ed.), Atlas of Neotropical Lepidoptera. Volume 5A. Association for Tropical Lepidoptera, Scientific Publishers, Gainesville.

Linck, R. W. (1982) The structure of microtubules. *Ann. N. Y. Acad. Sci.*, **383**: 98-121.

Linck, R. W. (1990) Tektins and microtubules. *Advances in Cell Biology*, **3**: 35-65.

Linck, R. W., & Langevin, G. L. (1982) Structure and chemical composition of insoluble filamentous components of sperm flagella microtubules. *J. Cell Sci.*, **58**: 1-22.

Linck, R. W., Amos, L. A. & Amos, B. (1985) Localisation of Tektin filaments in microtubules of sea urchin sperm flagella by immunoelectron microscopy. *J. Cell Biol.*, **100**: 126-135.

Linck, R. W., Albertini, D. F., Kenney, D. M. & Langevin, G. L. (1982) Tektin filaments: chemically unique filaments of sperm flagellar microtubules. *Prog. Clin. Biol. Res.*, **80**: 127-132.

Machado, C. A., Kliman, R. M., Markert, J. A. & Hey, J. (2002) Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and its close relatives. *Mol. Biol. Evol.*, **19**: 472-488.

Maddison, W. P., & Maddison, D. R. (1997) MacClade: Analysis of phylogeny and character evolution. Sinauer Associated, Sunderland. Massachussets.

Mallet, J. (2005) Hybridization as an invasion of the genome. *Trends Ecol. Evol.*, **20** (5) 229-237.

Nei, M. (1987) Molecular Evolutionary Genetics. Columbia University Press, New York.

Neigel, J. E. & Avise, J. C. (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation, pp. 515-534. In S. Karlin and E. Nevo (eds.), Evolutionary Processes and Theory. Academic Press, New York.

Nojima, D., Linck, R. W., & Egelman, E. H. (1995) At least one of the protofilaments in flagella microtubules is not composed of tubulin. *Curr. Biol.*, **5**: 158-167.

Norrander, J. M., Amos, L. A. & Linck, R. W. (1992) Primary structure of tekin A1: Comparison with intermediate filament proteins and a model for its association with tubulin. *Proc. Natl. Acad. Sci. U. S. A.*, **89**: 8567-8571.

Norrander, J. M., Perrone, C. A., Amos, L. A. & Linck, R. W. (1996) Structural Comparison of Tektins and Evidence for Their Determination of Complex Spacings in Flagellar Microtubules. *J. Mol. Biol.*, **257**: 385-397.

Ota, A., Kusakabe, T., Sugimoto, Y., Takahashi, M., Nakajima, Y., Kawaguchi, Y. & Koga, K. (2002) Cloning and characterisation of testis-specific Tektin in *Bombyx mori*. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.*, **133**: 371-382.

Pamilo, P. & Nei, M. (1988) Relationships between gene trees and species trees. *Mol. Biol. Evol.*, **5**: 568-583.

Philippe, H. & Douady, C. J. (2003) Horizontal gene transfer and phylogenetics. *Curr. Opin. Microbiol.*, **6**: 498-505.

Pirner, M. A. & Linck, R. W. (1994) Tektins are heterodimeric polymers in flagellas microtubules with axial periodicities matching the tubulin lattice. *J. Biol. Chem.*, **269**: 31800-31806.

Rokas, A., Williams, B. L., King, N. & Carroll, S. B. (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature (Lond.)*, **425**: 798-804.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., & Flook, P. (1994) Evolution, weighting, and phylogenetic usefulness of mitochondrial genes with a compilation of conserved PCR primers. *Ann. Entomol. Soc. Am.*, **87**: 651-701.

Sorenson, M. D. (1999) TreeRot, version 2. Boston University, Boston, Massachusetts.

Swofford, D. L. (2000) *PAUP**: Phylogenetic Analysis Using Parsimony (* and Other Methods). Sinauer Associates, Sunderland, Massachusetts.

Syvanen, M. (1994) Horizontal gene transfer: evidence and possible consequences. *Ann. Rev. Genet.*, **28**: 237-261.

Takahata, N., & Nei, M. (1985) Gene genealogy and variance of interpopulational nucleotide differences. *Genetics*, **110**: 325-344.

CHAPTER FIVE

MITOCHONDRIAL DNA PROVIDES AN INSIGHT INTO THE MECHANISMS DRIVING DIVERSIFICATION IN THE ITHOMIINE BUTTERFLY *HYPOSCADA ANCHIALA* (LEPIDOPTERA: NYMPHALIDAE, ITHOMIINAE)

Abstract

Geographic subspecies of several ithomiine butterflies on the lower east Andean slopes display a black and orange "melanic tiger" aposematic wing pattern that occurs from Colombia to Bolivia, while geographically adjacent lowland subspecies typically bear a coloured, "tiger" pattern. However, it is not clear whether subspecies with similar wing patterns in different regions have arisen through independent events of convergent adaptation, possibly through parapatric differentiation, or result from allopatric differentiation, as proposed by the refuge hypothesis. Here, we examine geographic patterns of divergence in the widespread and common ithomiine butterfly *Hyposcada anchiala.* We present phylogenetic hypotheses for 5 subspecies of *H. anchiala*, based on 1567 bp mitochondrial DNA. All topologies indicated that a single switch in mimetic pattern best explained the wing patterning of the *H. anchiala* studied here. This finding suggests that the subspecies of *H. anchiala* studied here result from at least two stages of differentiation, and is consistent with a single colonisation into a novel altitudinal zone coincident with a wing pattern switch, followed by subsequent divergence within, rather than across altitudinal zones. The subspecies divergences indicated diversifications were consistent with occurring during the Pleistocene. Furthermore, the lowland subspecies were more recently derived than the montane taxa, in contrast to predictions of the "Andean species pump" hypothesis.

Introduction

Butterfly species richness reaches a global peak at the western edge of the Amazon basin and adjacent east Andean foothills (Robbins & Opler 1996). To date, however, there is little agreement on the main causes for this high diversity, and a number of competing hypotheses have been proposed. The Pleistocene forest refuge hypothesis (Haffer 1969) has perhaps received the most attention (Brown 1979, Brown 1982, Sheppard et al. 1985, Brown 1987, Hall & Harvey 2002), and indeed a number of east Andean lower elevation putative forest refugia have been identified (Brown 1979, Brown 1982). However, critics argue that parapatric differentiation across environmental gradients could produce similar biogeographic patterns (Benson 1982, Endler 1982). In particular, the rapid changes in both biotic and abiotic environments at the base of the eastern Andes could have helped generate the high species diversity in this region. Fjeldså (1994) went on to propose that Andean slopes act as a "species pump" for birds, where new species originate before subsequent dispersal into the Amazon lowlands.

The first studies of the effects of putative forest refugia on neotropical butterflies were based principally on two nymphalid subfamilies, the Heliconiinae and the Ithomiinae (Brown 1982, Brown 1987). Both groups are notable for their members being unpalatable to predators and aposematically coloured, with often remarkable geographic differentiation in warning colour pattern. Heliconiines and ithomiines are extensively involved in Müllerian mimicry, in which two or more distasteful species converge on a common warning signal which is more easily learnt and remembered by predators (Müller 1879). Theory predicts that Müllerian mimicry increases in value as the number of individuals displaying the warning signal increases, due to a lower rate of attack per individual.

Brown (1982) suggested that such species were likely to differentiate in isolated forest refugia because they are relatively sedentary and because random fluctuations in abundance of co-mimics could drive mimicry pattern change (Brown & Benson 1974, Joron et al. 1999). However, mimetic species are also expected to be highly susceptible to parapatric differentiation. Mimetic colour patterns also serve as mating cues (Jiggins et al. 2001), so shifts in mimicry pattern, which might be driven by geographic variation in abundance of co-

mimics, are likely to disrupt gene flow. In the east Andean foothills mimetic butterfly faunas change rapidly over elevations of several hundred metres and distances of only a few kilometres (Keith Willmott pers. obs.), so the potential for colonisation of new mimetic environments is high.

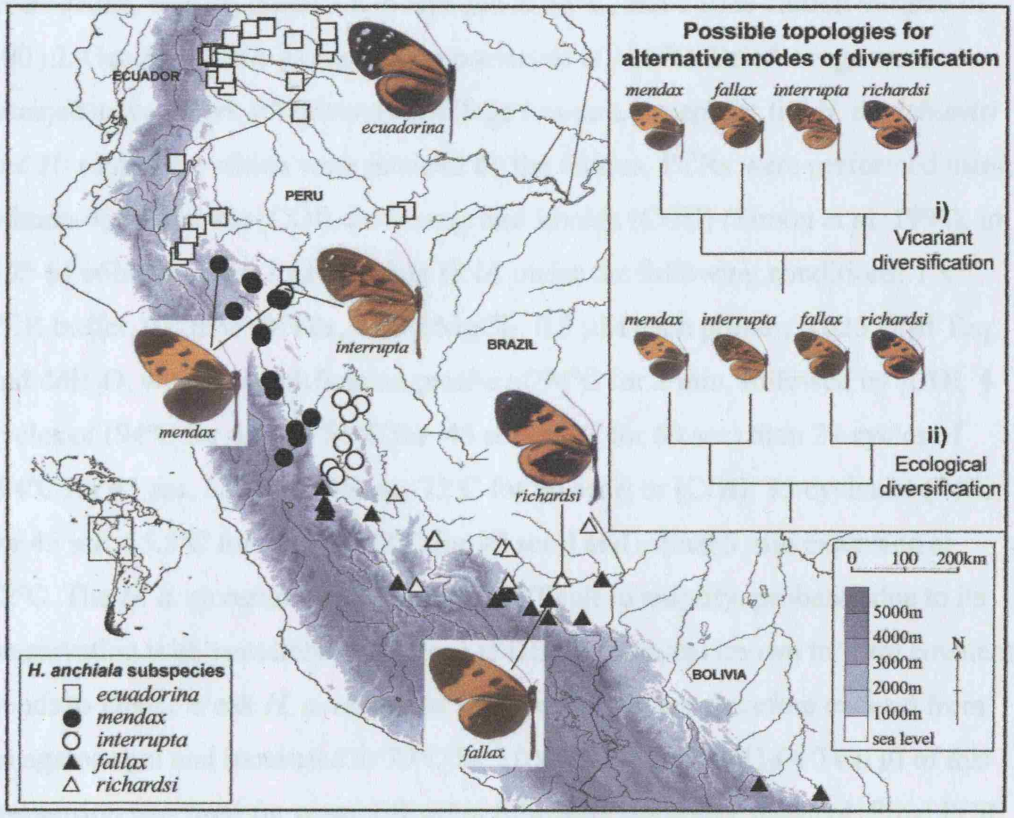
This debate still remains today partly because of the difficulties in testing between vicariance or ecological adaptation as causes of divergence. Here, we examine the possible roles of these two causes in explaining divergence in a single west Amazonian ithomiine species, *Hyposcada anchiala*. The Ithomiinae (Lepidoptera: Nymphalidae) is an exclusively Neotropical subfamily which includes ~355 species (Lamas 2004), all of which are thought to be highly unpalatable. These butterflies form part of multiple, diverse mimicry rings in a single location (Beccaloni 1997a, Beccaloni 1997b, Joron & Mallet 1998). *Hyposcada anchiala* occurs from eastern Panama to Bolivia and western Brazil, and contains 12 recognised subspecies (Lamas 2004). These subspecies display wing patterns of two distinct mimicry complexes, the orange and black tiger (melanic tiger) complex and the tiger complex. These complexes are strongly altitudinally zoned, constrained to high altitude (submontane) and low altitude sites respectively. In this paper we focus on relationships between the submontane melanic tigers, *H. a. mendax* Fox from N. Peru and *H. a. fallax* (Staudinger) from S. Peru and Bolivia, and the geographically adjacent, lowland tiger-pattern subspecies, *H. a. interrupta* Tessmann in the north and *H. a. richardsi* Fox, in the south (Fig. 1.).

If individuals cluster by subspecies, there are 26 possible rooted cladograms for *H. a. mendax*, *H. a. fallax*, *H. a. interrupta* and *H. a. richardsi*. Of those, 6 are not informative about the associated mimetic-altitudinal shifts, for example an unresolved quadchotomy, 7 support just one mimetic-altitudinal shift, and 13 support two mimetic-altitudinal shifts. An example of the 7 topologies which support just a single shift in mimicry and altitude is a bifurcating topology, with both montane tiger melanic subspecies in one clade, and both non-melanic tiger, lowland subspecies in the other clade (as depicted in Fig. 1), this can best be explained by a vicariant diversification of two previously widespread ancestral populations. The other 13 topologies support two shifts of mimicry and altitude, and thus ecological diversification (Benson 1982, Endler 1982), an example is a sister pairing of the two subspecies from N.

Peru (one montane, melanic and one lowland, non-melanic tiger), and a sister pairing of the two subspecies from S. Peru (one montane, melanic and one lowland, non-melanic tiger) (as depicted in Fig. 1.). Under this scenario, local processes drove the diversification of an ancestral N. Peru *H. anchiala* into two northern subspecies, a process paralleled independently in S. Peru.

Here, we use 1567 bp mitochondrial DNA (mtDNA) to reconstruct relationships within *H. anchiala* and examine whether resulting phylogeographic patterns are more consistent with expectations of geographic (vicariant) or ecological diversification.

Figure 1. Map showing the butterfly collection localities for *H. a. mendax*, *H. a. fallax*, *H. a. interrupta*, *H. a richardsi* and *H. a. ecuadorina* (subspecies symbols are as in later phylogentic hypotheses, see Fig. 3 and Fig. 4). The lines connecting figured butterflies indicate sampling locations for the molecular samples. Inset shows 2 of the 26 possible rooted topologies for *H. a. mendax*, *H. a. fallax*, *H. a. interrupta* and *H. a richardsi*. The displayed topologies are consistent with i) vicariant, and ii) ecological, diversification theories.



Materials and methods

Genomic DNA was analysed from the four described *H. anchiala* subspecies, plus an additional non-melanic tiger pattern subspecies, *H. a. ecuadorina* Bryk. Outgroup taxa were chosen to represent two closely related species, *H. virginiana* (Hewitson) and *H. zarepha* (Hewitson), as determined by the sequencing of five *Hyposcada* species with two, independent nuclear regions, *wingless* and *elongation factor 1-α* (data not shown) as well as mtDNA data (Table 1).

A specimen of *H. virginiana* in 20% Dimethylsulphoxide, 0.25M EDTA and saturated NaCl solution (DMSO solution) was identified and provided by Chris Jiggins (University of Edinburgh) and a dried specimen of *H. a. richardsi* was provided by Gerardo Lamas. All other specimens were collected by the authors, preserved in DMSO solution, and identified to subspecies by Gerardo Lamas and Keith Willmott. DNA was extracted from one third of the abdomen (*H. a. richardsi*) or one third of the thorax (all other specimens) using the DNAeasy kit (QIAgen, West Sussex, UK), according to the manufacturer's instructions, with an initial 3 h incubation at 55°C, and a final elution volume of 300 µl. Genomic DNA extracts were preserved at –20°C. Dried wings were retained as vouchers at University College London, except for the *H. a. richardsi* and *H. virginiana* which were retained by the donors. PCRs were performed using primers Jerry and Pat (COI), or George and Imelda (COII) (Simon et al. 1994), in a 25 µl volume, using 2 µl template DNA under the following conditions: 1 x PCR buffer, 0.6 mM dNTPs, 4 mM $MgCl_2$, 0.5 µM each primer, 0.025 U/µl Taq, and ddH_2O, with an amplification profile of 94°C for 2 min, followed by [COI: 4 cycles of (94°C for 45 sec, 51°C for 45 sec, 72°C for 60 sec) then 29 cycles of (94°C for 45 sec, 51°C for 45 sec, 72°C for 90 sec)] or [COII: 33 cycles of (94°C for 45 sec, 55.5°C for 45 sec, 72°C for 90 sec)] and a final 5 min extension at 72°C. The *H. a. richardsi* sample proved difficult to amplify, probably due to its preservation with paradichlorobenzene (PDB), a chemical known to form covalent bonds to DNA. Weak *H. a. richardsi* PCR products were therefore excised from an agarose gel and incubated at 70°C for 10 min in 1.5 ml ddH_2O. Two µl of this suspension was used for reamplification following the above protocol. Final PCR products were purified using the QIAquick PCR purification kit (QIAgen, West

Table 1. Specimen information, collection localities and respective Genbank accession numbers for the COI-tRNA-COII region amplified in the 13 individuals studied here. Submitted data cover the entire length of the region amplified (sequence lengths varied between 1567 and 1611 bp due to loss of data at chromatogram ends).

Taxon	Voucher number	Collection locality	Approximate collection altitude (m)	Genbank Accession
H. anchiala interrupta	02-512	PERU: San Martín: km 7.2 Pongo-Barranquita. 06°17'20"S, 76°13'41"W	150	DQ078355
H. anchiala interrupta	02-2105	PERU: San Martín: km 7.2 Pongo-Barranquita. 06°17'20"S, 76°13'41"W	150	DQ078356
H. anchiala interrupta	02-1293	PERU: San Martín: km 7.2 Pongo-Barranquita. 06°17'20"S, 76°13'41"W	150	DQ078312
H. anchiala mendax	02-1644	PERU: San Martín: Puente Serranoyacu. 05°40'S, 77°40'W	1150	DQ078357
H. anchiala mendax	02-1645	PERU: San Martín: Puente Serranoyacu. 05°40'S, 77°40'W	1150	DQ078358
H. anchiala mendax	02-716	PERU: San Martín: Puente Serranoyacu. 05°40'S, 77°40'W	1150	DQ078359
H. anchiala mendax	02-1602	PERU: San Martín: La Antena. 06°27'18"S, 76°17'54"W	800-1000	DQ078360
H. anchiala mendax	02-2141	PERU: San Martín: La Antena. 06°27'18"S, 76°17'54"W	800-1000	DQ078361
H. anchiala fallax	02-3519	PERU: Madre de Dios: Mazuko. 13°06'S, 70°22'W	350	DQ078477
H. anchiala ecuadorina	Ec 452	ECUADOR: Napo: Jatun Sacha. 01°04'S, 77°36'W	450	DQ078474
H. anchiala richardsi	G1	PERU: Madre de Dios: Río Los Amigos. 12°35'S, 7005'W	270	DQ078475
H. virginiana	8283	PANAMA: Cerro Campana. 08°68'74"N, 79°91'97"W	900-1000	DQ078476
H. zarepha flexibilis	02-197	PERU: San Martín: Chumía. 06° 36'57"S, 76° 11'10"W	300-450	DQ078362

Table 2. Absolute pairwise distances for all *H. anchiala* individuals (below diagonal), mean Hasegawa-Kishino-Yano (HKY85) subspecies pairwise distances (in %, above diagonal) and mean estimated time since divergence of subspecies (in years (y), to the nearest 5,000, above diagonal).

Hyposcada anchiala subspecies and specimen number		*interrupta* (n = 3)			*mendax* (n = 5)					*fallax* (n = 1)	*ecuadorina* (n = 1)	*richardsi* (n = 1)
		02-512	02-2105	02-1293	02-1644	02-1645	02-716	02-1602	02-2141	02-3519	Ec 452	G1
interrupta	02-512		X	X	240,000 y 0.55%					140,000 y 0.32%	55,000 y 0.13%	85,000 y 0.19%
	02-2105	0		X								
	02-1293	0	0									
mendax	02-1644	9	9	9		X	X	X	X	310,000 y 0.72%	285,000 y 0.66%	320,000 y 0.74%
	02-1645	9	9	9	0		X	X	X			
	02-716	7	7	7	6	6		X	X			
	02-1602	9	9	9	2	2	6		X			
	02-2141	9	9	9	0	0	6	2				
fallax	02-3519	5	5	5	12	12	8	12	12		195,000 y 0.45%	220,000 y 0.51%
ecuadorina	Ec 452	2	2	2	11	11	9	9	11	7		140,000 y 0.32%
richardsi	G1	3	3	3	12	12	10	12	12	8	5	

Sussex, UK), according to the manufacturer's protocol and sent to a commercial facility for cycle sequencing using the PCR primers, precipitation and sequencing. Sequences were edited using SeqEd v1.0.3 (Applied Biosystems, Inc). PAUP version 4.0b 10 (Swofford 2000) was used to calculate the numbers of variable and parsimony informative sites, and absolute and Hasegawa-Kishino-Yano (HKY85) pairwise divergences (Table 2). The constancy of evolutionary rates across the tree was tested in PAUP with a likelihood ratio test, and a molecular clock calibrated as 2.3% sequence divergence per million years (Brower 1994) was used to estimate divergence dates (Table 2). For comparison, Neighbour Joining (NJ), Maximum Parsimony (MP) and Maximum Likelihood (ML) analyses were performed in PAUP, and a Bayesian hypothesis was generated using MrBayes 3.0 (Huelsenbeck & Ronquist 2001), with *H. virginiana*, or *H. zarepha*, or (*H. zarepha* + *H. virginiana*) as the outgroup taxa. NJ trees were generated using HKY85 distances with random break ties, confidence in each node was assessed by performing 1000 rounds of NJ bootstrapping. MP topologies were conducted using a 10 replicate heuristic search with tree-bisection-reconnection (TBR) branch swapping, with random stepwise addition of sequences, support for clades was assessed by 1000 rounds of full heuristic bootstrapping with TBR branch swapping. ML trees were generated starting from random trees, using a heuristic search with TBR, random sequence addition, with the best fit model of sequence evolution (HKY+I+G, by hLRT) and parameter estimates as identified by Modeltest 3.04 (Posada & Crandall 1998). Bayesian analyses were performed with parameters identified by Modeltest 3.04 (nst=2, rates=invgamma), with four simultaneous chains run for 1,000,000 generations, sampling a tree every 100 generations. Consensus trees with branch support in the form of posterior probabilities were derived from the final 9,000 trees (representing the final 900,000 generations), after confirmation that likelihood values had stabilised after the first 100,000 generations. All possible unrooted trees were generated and corresponding minimum evolution (ME) scores calculated, for one representative of each *H. anchiala* subspecies plus a *H. virginiana* outgroup, using distance settings.

Results

Sequences up to 1611 bp in length were successfully obtained from 13 individuals, representing five *H. anchiala* subspecies and two outgroup species (DQ078312, DQ078355-62, and DQ078474-77). To avoid problems associated with missing characters, sequences were edited to a final alignment covering the 1567 bp region amplified in all specimens. The 1567 bp span 805 bp COI (alignment positions 1-805), 62 bp leucine-tRNA (806-867) and 700 bp COII (868-1567). Excluding the outgroup taxa, just 19 (1.21%) sites were variable, 10 in COI and nine in COII, of which 11 (0.70%) were parsimony informative. Of these 19 polymorphic sites, 16 differed in just one species (4 fixed and 5 polymorphic in *H. a. mendax*, 3 in *H. a. fallax*, 3 in *H. a. richardsi*, and 1 in *H. a. ecuadorina*), and three polymorphisms were shared; C in *H. a. mendax* + *fallax*, and T in *H. a. ecuadorina* + *richardsi* + *interrupta*; G in *H. a. fallax* + *richardsi* + *interrupta*, A↔G in *H. a. mendax*, and A in *H. a. ecuadorina*; and A in *H. a. ecuadorina* + *richardsi* + *interrupta*, A↔T in *H. a. mendax*, and T in *H. a. fallax* (see Fig. 2).

Absolute *H. anchiala* pairwise divergences (Table 2) ranged from 0, within the *H. a. interrupta* or between *H. a. mendax* 02-1644, 02-1645 and 02-2141, to 12 between *H. a. ecuadorina* or *H. a. richardsi* and *H. a. mendax* 02-1644, 02-1645, 02-2141 and 02-1602. There was no significant rate heterogeneity among lineages of the clock enforced [ln likelihood = -2487.286] and non-clock enforced [ln likelihood= 2482.988] trees, suggesting the data conform to a molecular clock. Mean between subspecies divergence times were estimated using a mitochondrial clock (Brower 1994) to between 55,000 and 320,000 years ago, for (*H. a. interrupta* + *H. a. ecuadorina*) and (*H. a. mendax* + *H. a. richardsi*) respectively.

Overall, the *H. anchiala* phylogenies had short internal branches and thus were quite poorly resolved. The clearest case of phylogenetic agreement between all reconstructions was the well supported recovery of the five *H. a. mendax* specimens as a monophyletic clade, for example, in analyses using a *H. virginiana* + *H. zarepha* outgroup by bootstrap values of 94 (NJ) and 92 (MP), and a Bayesian posterior probability of 0.72. Although the arrangements of *H. anchiala* differed both with outgroup used, and reconstruction method, the topologies could be categorised into three, repeatedly generated types. NJ and

Figure 2. Sequence alignment of the full 1611 base pairs (bp) obtained from all individuals of the 5 *H. anchiala* subspecies studied here, showing the COI (upper case, black), tRNA (lower case, boxed), and COII (upper case, blue), gene regions. Subspecies and specimen numbers are as follows: *H. a. interrupta*, 02-512, 02-2105, 02-1293; *H. a. ecuadorina*, Ec 452; *H. a. richardsi*, G1; *H. a. mendax*, 02-1644, 02-1645, 02-2141, 02-1602, 02-716; and *H. a. fallax, 02-3519*. Positions of the 19 variable sites (*) are indicated along the bottom line of the alignment.

```
bp 1-78:
02-512      --ATCCAGAAGTTTATATTTTAATTCTTCCAGGATTTGGAATAATCTCTCATATTATTTCCCAAGAAAGAGGAAAAAA
02-2105     --ATCCAGAAGTTTATATTTTAATTCTTCCAGGATTTGGAATAATCTCTCATATTATTTCCCAAGAAAGAGGAAAAAA
02-1293     TCATCCAGAAGTTTATATTTTAATTCTTCCAGGATTTGGAATAATCTCTCATATTATTTCCCAAGAAAGAGGAAAAAA
Ec 452      ------------------------TCTTCCAGGATTTGGAATAATCTCTCATATTATTTCCCAAGAAAGAGGAAAAAA
G1          TCATCCAGAAGTTTATATTTTAATTCTTCCAGGATTTGGAATAATCTCTCATATTATTTCCCAAGAAAGAGGAAAAAA
02-1644     TCATCCAGAAGTTTATATTTTAATTCTTCCAGGGTTTGGAATAATCTCTCATATTATTTCTCAAGAAAGAGGAAAAAA
02-1645     TCATCCAGAAGTTTATATTTTAATTCTTCCAGGGTTTGGAATAATCTCTCATATTATTTCTCAAGAAAGAGGAAAAAA
02-2141     -----CAGAAGTTTATATTTTAATTCTTCCAGGGTTTGGAATAATCTCTCATATTATTTCTCAAGAAAGAGGAAAAAA
02-1602     ----CCAGAAGTTTATATTTTAATTCTTCCAGGGTTTGGAATAATCTCTCATATTATTTCTCAAGAAAGAGGAAAAAA
02-716      -CATCCAGAAGTTTATATTTTAATTCTTCCAGGGTTTGGAATAATCTCTCATATTATTTCTCAAGAAAGAGGAAAAAA
02-3519     -------GAAGTTTATATTTTAATTCTTCCAGGATTTGGAATAATCTCTCATATTATTTCCCAAGAAAGAGGAAAAAA
            ---------------------------------*------------------------*-----------------

bp 79-156:
02-512      AGAAACTTTTGGATCTTTAGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCACCA
02-2105     AGAAACTTTTGGATCTTTAGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCACCA
02-1293     AGAAACTTTTGGATCTTTAGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCACCA
Ec 452      AGAAACTTTTGGATCTTTAGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCACCA
G1          AGAAACTTTTGGATCTTTAGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCACCA
02-1644     AGAAACTTTTGGATCTTTGGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCATCA
02-1645     AGAAACTTTTGGATCTTTGGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCATCA
02-2141     AGAAACTTTTGGATCTTTGGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCATCA
02-1602     AGAAACTTTTGGATCTTTGGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCATCA
02-716      AGAAACTTTTGGATCTTTAGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATTGTATGAGCTCATCA
02-3519     AGAAACTTTTGGATCTTTAGGAATAATTTATGCTATAATAGCAATTGGATTATTAGGATTTATCGTATGAGCTCACCA
            ------------------*-----------------------------------------------*----------*--

bp 157-234:
02-512      TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACTGG
02-2105     TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACTGG
02-1293     TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACTGG
Ec 452      TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACTGG
G1          TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACTGG
02-1644     TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACCGG
02-1645     TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACCGG
02-2141     TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACCGG
02-1602     TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACCGG
02-716      TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACTGG
02-3519     TATATTTACTGTAGGTATAGATATTGATACTCGTGCTTATTTTACTTCTGCTACAATAATTATTGCCGTTCCAACTGG
            ------------------------------------------------------------------------------*--

bp 235-312:
02-512      AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
02-2105     AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
02-1293     AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
Ec 452      AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
G1          AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
02-1644     AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
02-1645     AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
02-2141     AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
02-1602     AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
02-716      AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
02-3519     AATTAAAATTTTTAGTTGATTAGCTACTTTACATGGTACTCAAATTAATTATAGACCTTCAATTTTATGAAGATTAGG
            ------------------------------------------------------------------------------
```

Figure 2 (continued)

```
bp 313-390:
02-512      ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
02-2105     ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
02-1293     ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
Ec 452      ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
Gl          ATTTGTATTTTTATTTACTGTTGGGGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
02-1644     ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
02-1645     ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
02-2141     ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
02-1602     ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
02-716      ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
02-3519     ATTTGTATTTTTATTTACTGTTGGAGGATTAACAGGAGTAATTTTAGCTAATTCTTCAATCGATATTACTCTTCATGA
            ------------------------*-----------------------------------------------------

bp 391-468:
02-512      TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
02-2105     TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
02-1293     TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
Ec 452      TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGAGGATTTATTCA
Gl          TACATATTATGTAGTTGCTCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
02-1644     TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
02-1645     TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
02-2141     TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
02-1602     TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
02-716      TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
02-3519     TACATATTATGTAGTTGCCCATTTTCATTATGTACTTTCAATAGGAGCAGTATTTGCAATTTTAGGGGGATTTATTCA
            ------------------*-----------------------------------------------*-----------

bp 468-546:
02-512      TTGATATCCTTTATTTACTGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
02-2105     TTGATATCCTTTATTTACTGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
02-1293     TTGATATCCTTTATTTACTGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
Ec 452      TTGATATCCTTTATTTACTGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
Gl          TTGATATCCTTTATTTACTGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
02-1644     TTGATATCCTTTATTTACCGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
02-1645     TTGATATCCTTTATTTACCGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
02-2141     TTGATATCCTTTATTTACCGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
02-1602     TTGATATCCTTTATTTACCGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
02-716      TTGATATCCTTTATTTACCGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
02-3519     TTGATATCCTTTATTTACCGGATTAACTTTAAATCCTTATTTATTAAAAATTCAATTTATTACTATATTTTTAGGGGT
            ------------------*-----------------------------------------------------------

bp 547-624:
02-512      TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
02-2105     TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
02-1293     TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
Ec 452      TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
Gl          TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
02-1644     TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
02-1645     TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
02-2141     TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
02-1602     TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
02-716      TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
02-3519     TAATCTTACTTTCTTTCCTCAACATTTTTTAGGTCTAGCTGGAATACCTCGTCGTTATTCAGATTATCCAGATAGATT
            ------------------------------------------------------------------------------

bp 625-702:
02-512      TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
02-2105     TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
02-1293     TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
Ec 452      TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
Gl          TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
02-1644     TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
02-1645     TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
02-2141     TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
02-1602     TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
02-716      TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
02-3519     TGTTTCTTGAAATATTATTTCTTCTTTCGGATCATATATTTCATTTTTATCAATAATTATAATTATTATTATTATTTG
            ------------------------------------------------------------------------------
```

Figure 2 (continued)

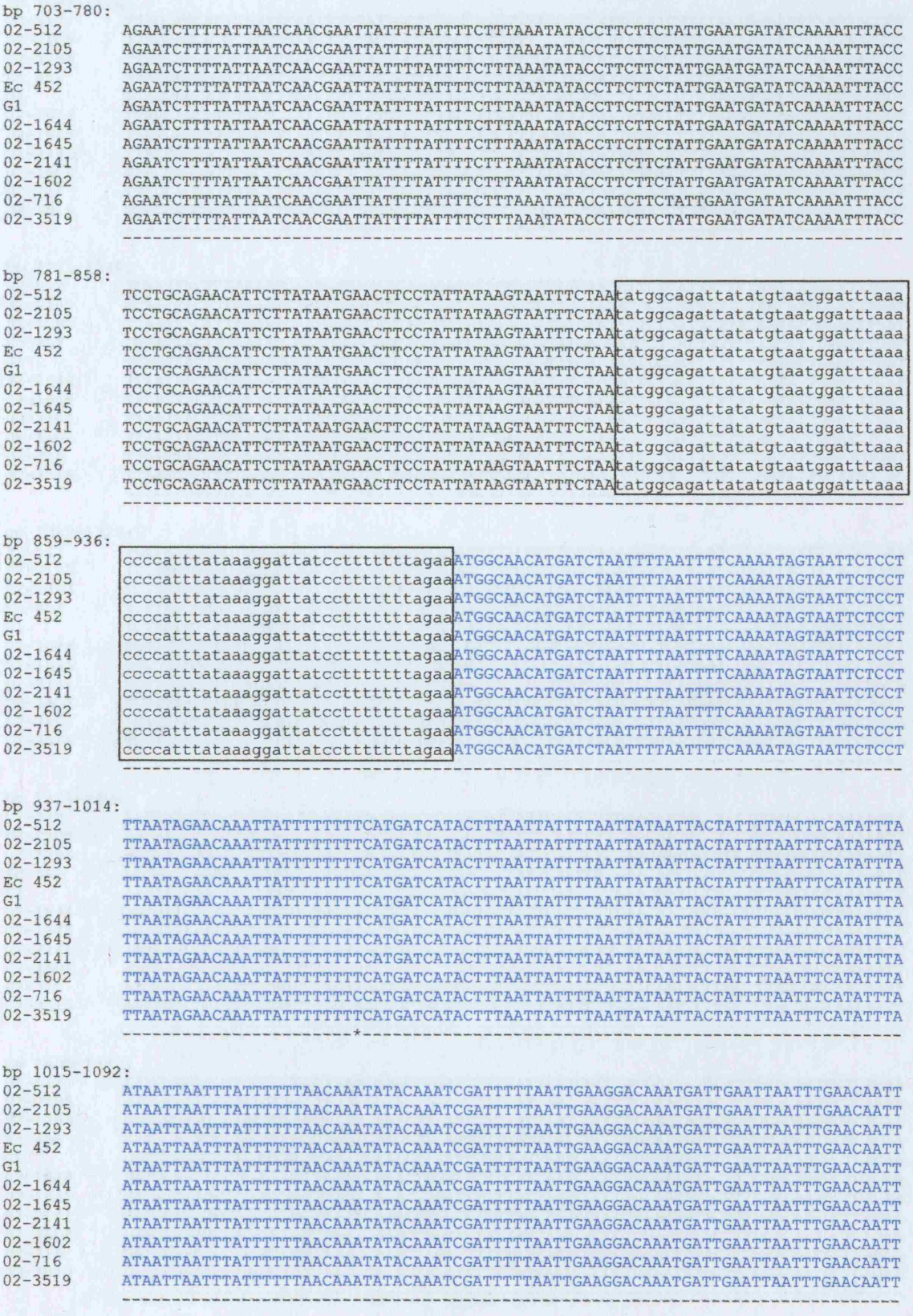


Figure 2 (continued)

```
bp 1093-1170:
02-512      TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTACGATTATTATATCTTTTAGATGAAATTAATAACCCT
02-2105     TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTACGATTATTATATCTTTTAGATGAAATTAATAACCCT
02-1293     TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTACGATTATTATATCTTTTAGATGAAATTAATAACCCT
Ec 452      TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTACGATTATTATATCTTTTAGATGAAATTAATAACCCT
G1          TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTCTACGATTATTATATCTTTTAGATGAAATTAATAACCCT
02-1644     TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTGCGATTATTATATCTTTTAGATGAAATTAATAACCCT
02-1645     TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTGCGATTATTATATCTTTTAGATGAAATTAATAACCCT
02-2141     TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTGCGATTATTATATCTTTTAGATGAAATTAATAACCCT
02-1602     TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTACGATTATTATATCTTTTAGATGAAATTAATAACCCT
02-716      TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTACGATTATTATATCTTTTAGATGAAATTAATAACCCT
02-3519     TTACCAGCATTTACTTTAATTTTTATTGCTTTACCTTCTTTACGATTATTATATCTTTTAGATGAAATTAATAACCCT
            ----------------------------------------*-*-----------------------------------

bp 1171-1248:
02-512      TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
02-2105     TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
02-1293     TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
Ec 452      TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
G1          TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
02-1644     TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
02-1645     TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
02-2141     TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
02-1602     TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
02-716      TTAATTACTTTAAAATCAATCGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
02-3519     TTAATTACTTTAAAATCAATTGGACATCAATGATATTGAAGATATGAATATTCTGATTTTTTTAATATTGAATTTGAT
            --------------------*-------------------------------------------------------

bp 1249-1326:
02-512      TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
02-2105     TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
02-1293     TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
Ec 452      TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
G1          TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
02-1644     TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
02-1645     TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
02-2141     TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
02-1602     TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
02-716      TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGCTTATTAGATGTTGATAATCGAATTATTTTACCTATA
02-3519     TCTTATATAACTCAATCATCTGAATATATAAATAATTTTCGTTTATTAGATGTTGATAATCGAATTATTTTACCTATA
            ----------------------------------------*-----------------------------------

bp 1327-1404:
02-512      AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGGGTAAAAATT
02-2105     AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGGGTAAAAATT
02-1293     AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGGGTAAAAATT
Ec 452      AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGAGTAAAAATT
G1          AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGGGTAAAAATT
02-1644     AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGGGTAAAAATT
02-1645     AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGGGTAAAAATT
02-2141     AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGGGTAAAAATT
02-1602     AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGAGTAAAAATT
02-716      AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGGGTAAAAATT
02-3519     AATAATCAAATTCGTATTTTAGTAACTGCTACTGATGTTATTCATTCTTGAACAATTCCTGCTTTAGGGGTAAAAATT
            ------------------------------------------------------------------------*---------

bp 1405-1482:
02-512      GATGCAAATCCAGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
02-2105     GATGCAAATCCAGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
02-1293     GATGCAAATCCAGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
Ec 452      GATGCAAATCCAGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
G1          GATGCAAATCCAGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
02-1644     GATGCAAATCCAGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
02-1645     GATGCAAATCCAGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
02-2141     GATGCAAATCCAGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
02-1602     GATGCAAATCCAGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
02-716      GATGCAAATCCTGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
02-3519     GATGCAAATCCTGGTCGATTAAATCAAACTAGATTTTTTATTAATCGCCCTGGAATATACTTTGGACAATGCTCAGAA
            ----------*-------------------------------------------------------------------
```

Figure 2 (continued)

```
bp 1483-1560:
02-512      ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAATTAAT
02-2105     ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAATTAAT
02-1293     ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAATTAAT
Ec 452      ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAATTAAT
G1          ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAATTAAT
02-1644     ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAGTTAAT
02-1645     ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAGTTAAT
02-2141     ATTTGTGGA?CTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAGTTAAT
02-1602     ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAGTTAAT
02-716      ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAATTAAT
02-3519     ATTTGTGGAGCTAATCATAGTTTTATACCTATTGTAATTGAAAGAATTTCAATAATAAATTTTATTAAATGAATTAAT
            ------------------------------------------------------------------------*-----

bp 1561-1611:
02-512      AATTACTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA
02-2105     AATTACTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA
02-1293     AATTACTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA
Ec 452      AATTACTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA
G1          AATTACTCATTAGATGTCTGAAAGCAAGTAT--------------------
02-1644     AATTATTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA
02-1645     AATTATTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA
02-2141     AATTATTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA
02-1602     AATTATTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA
02-716      AATTATTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAAACCTTATTA
02-3519     AATTACTCATTAGATGTCTGAAAGCAAGTATAGGTCTCTTAA---------
            -----*-----------------------------------------------
```

MP analyses, using *H. zarepha* and (*H. zarepha* + *H. virginiana*) as the outgroup, shared a similar topology (as in Fig 3). All ML analyses and MP with a *H. virginiana* outgroup recovered the lowland *H. anchiala* as paraphyletic with regard to a sister pairing of *H. a. mendax* and *H. a. fallax*. The Bayesian analysis with a (*H. zarepha* + *H. virginiana*) outgroup shared a topology with NJ, MP and Bayesian analyses using just a *H. virginiana* outgroup (as in Fig 4). In addition, a fourth arrangement of a (*H. a. mendax* - *H. a. fallax*) and (*H. a. Richards i- H. a. interrupta- H. a. ecuadorina*) sister pairing was recovered once, by Bayesian analyses with a *H. zarepha* outgroup. In summary, of the four obtained topologies, two supported the lowland non-melanic *H. anchiala* having evolved from their montane melanic counterparts, one supported the vicariant hypotheses depicted in Fig. 1., and the other paired the montane melanic butterflies, but didn't indicate their relationship to the non-melanic lowland *H. anchiala*. Although these arrangements differed in the direction of the wing pattern switch and altitudinal shift, all four topology types support the main finding, that *H. anchiala* wing patterns can be most parsimoniously explained by a single switch in mimetic pattern and altitudinal range.

One hundred and five different trees were generated using distance settings, to represent all possible topologies. The mean ME score for those trees

most parsimoniously explained by one mimetic pattern shift was 0.0429 (n=21, range 0.0419-0.0445, median 0.0427) (the ME score of the topology which mirrored that recovered by NJ analysis was 0.0429), and the mean ME score for those topologies most parsimoniously explained by two shifts was 0.0439 (n=84, range 0.0424-0.0454, median 0.0439).

Discussion

Despite the rapid substitution rate of mtDNA compared to most nuclear regions, sequence results yielded only low levels of divergence among *H. anchiala* subspecies. Tree reconstruction was compromised both by sequence invariance and possibly by the outgroup taxa assuming a position rather distant from the ingroup. Although tree reconstruction method, and use of *H. virginiana*, *H. zarepha*, or (*H. zarepha* + *H. virginiana*) as the outgroup affected ingroup arrangements, all the phylogenetic hypotheses were most parsimoniously explained by a single colour shift between melanic and non-melanic taxa, thus the most recent radiations appear to be within altitudinal zones rather than between them. Taken along with the estimated subspecies divergence times, this finding is consistent with geographic diversification during the Pleistocene (11,000 - 1.8 Mya). This result of geographic divergence within an altitudinal zone parallels the findings by Willmott (2001) and Willmott et al. (2001) based on morphological phylogenies of *Hypanartia* Hübner (Nymphalinae) and *Adelpha* Hübner (Limenitidinae). In these genera speciation across elevational gradients occurred, but appeared relatively more difficult, infrequent and more typical of basal diversifications. The most common form of differentiation was speciation within or among montane regions. In contrast, in a recent study on *Ithomiola* Felder & Felder (Riodinidae), Hall (2006) found multiple instances of parapatric differentiation up the sides of mountains. In *Ithomiola*, montane species in different regions are most closely related to adjacent species occurring at lower elevations, indicating that speciation across an altitudinal gradient is the rule in these butterflies.

Figure 3. Phylogenetic hypothesis of the 5 *H. anchiala* subspecies, plus *H. zarepha* and *H. virginiana* ougroups, inferred with maximum parsimony. Maximum parsimony bootstrap support values are given along branches. Subspecies symbols are as in Fig. 1.

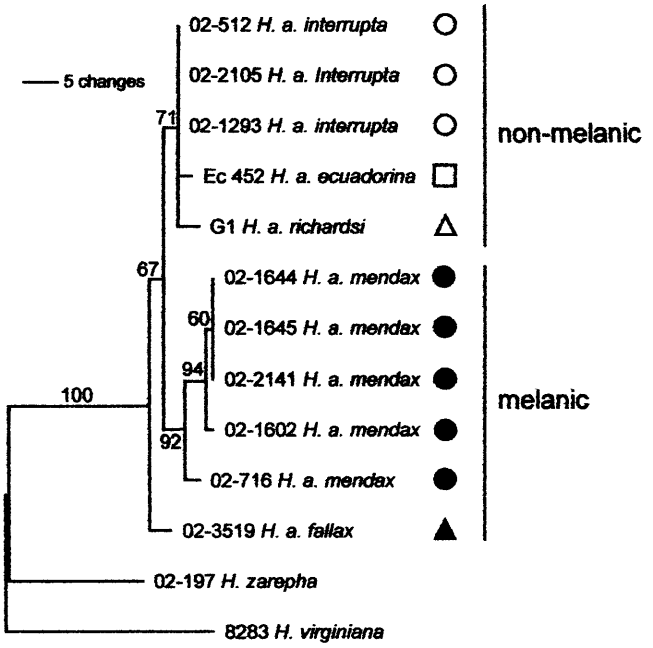
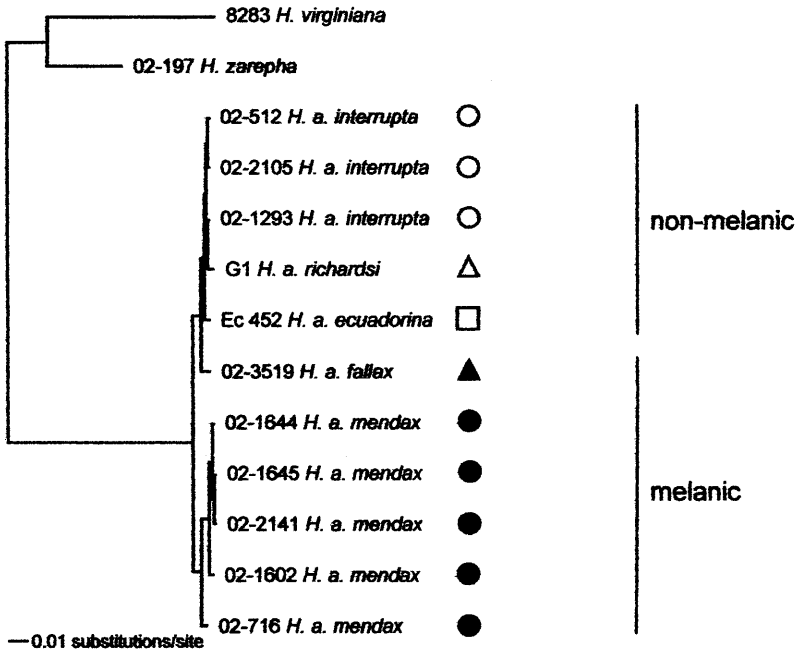


Figure 4. Phylogenetic hypothesis of the 5 *H. anchiala* subspecies, plus *H. zarepha* and *H. virginiana* ougroups, inferred with Bayesian methods. Subspecies symbols are as in Fig. 1. Only 3 clades have Bayesian probabilities over 0.5: (02-1644, 02-1645 and 02-2141), (02-1644, 02-1645, 02-2141 and 02-1602) and (02-1644, 02-1645, 02-2141, 02-1602 and 02-716) are supported with respective probabilities of 0.64, 0.85 and 0.72. Subspecies symbols are as in Fig. 1.



Although theories of Neotropical diversification have tended to concentrate on taxa recognised as species, studies based on subspecies are also of great value as these taxa may often form an early stage in the diversification continuum that eventually leads to complete speciation. The subspecies divergences here (Table 2) suggest that the lowland taxa are more recently derived (*H. a. richardsi* + *H. a. interrupta* 85,000 y) than the montane taxa (*H. a. mendax* + *H. a. fallax* 310,000 y). Our finding of older montane taxa appears to contrast Fjeldså's (1994) reports that Andean bird species were, on average, younger than Amazonian species. Fjeldså went on to suggest that the Andes acted as a species "pump", producing novelties that were absorbed into the Amazon, which acted as a species "museum". However, a problem lies in just looking at taxa categorised (maybe incorrectly) as distinct species, and ignoring diversity at other (often lower) levels. For example, if the taxa regarded as young Andean species by Fjeldså were actually old subspecies, and the Amazon contained many, well-differentiated young subspecies (not recognised as being differentiated in species level analysis), the conclusions regarding the location of diversifications would be the exact opposite to Fjeldså's.

However, the data and phylogenetic hypotheses presented here must be interpreted with caution. Only the *H. a. mendax* clade is repeatedly well supported. Additionally, the topologies represent the mtDNA tree, which is not necessarily equivalent to the species tree, due to processes including horizontal transfer, introgression and ancestral polymorphism (Brower et al. 1996). Given that the five *H. anchiala* subspecies here diverged moderately to rapidly on an evolutionary scale (within the last 300,000 years) the internodal branches are short, increasing the chance that ancestral copies failed to coalesce before diversification and therefore reducing the likelihood that mtDNA reflects the true organismal phylogeny (Maddison 1997). Additionally, there is evidence that introgressive gene flow may have considerably shaped the *H. anchiala* mtDNA topologies. Although distinguished by a mean pairwise divergence of 0.5% (corresponding to a radiation 220,000 years ago), intermediates are known between *H. a. fallax* and *H. a. richardsi* (Lamas, pers. obs.). One intermediate is also known between *H. a. fallax* and *H. a. interrupta*. That these subspecies are capable of interbreeding introduces potential for introgressive gene flow. Even

rare events can have a large impact, for example, an advantageous mutation could lead to a selective mtDNA sweep. Such gene flow can make two taxa appear to have diverged more recently than their actual original divergence time, creating a particular problem in studies, such as this, when the timing of the diversification event is critical for making inferences (Barraclough & Nee 2001).

To date, few studies have taken a phylogeographic approach to investigate the diversification of mimetic Neotropical butterflies. Despite 10 years since publication, Brower's (1994, 1996) mtDNA data for *Heliconius* butterflies remains the most significant contribution to this field. Our estimates of subspecies divergence dates are largely consistent with Brower's findings, but in contrast we found no evidence that similarly appearing mimetic races were mainly the result of convergent evolution. Brower (1994, 1996) found that the *Heliconius* mtDNA topology largely mirrored the biogeographic distributions, but that there was a lack of concordance between geography and wing pattern. The dissimilarity between *Hyposcada anchiala* and *Heliconius* might simply be an artefact of the chosen loci, and given the above concerns that mtDNA might not be a reliable foundation for subsequent inferences, it would be of value to corroborate both data sets with independent nuclear loci. Unfortunately such research is currently hampered as few molecular markers are available which evolve rapidly enough to allow reliable phylogenetic reconstruction of closely related groups of Lepidoptera. This study could also be improved by sequencing additional *H. anchiala* subspecies, but sufficiently preserved samples are typically hard to obtain. It would also be of interest to investigate other species with parallel upland-lowland distribution patterns. Notwithstanding these concerns, it seems likely that differences in ithomiine and heliconiine mimicry evolution reflect genuinely different evolutionary histories. Such differences have previously been reported, for example, in the ithomiine genus *Ithomia* Hübner mimicry shifts between species occur rarely (Jiggins et al. 2006), in sharp contrast to *Heliconius* (Brower 1996).

Although offering detailed insights into the mechanisms contributing to Neotropical diversification, it is likely that different taxa have experienced highly heterogeneous modes of diversification and thus individual examples are poor at elucidating the general patterns which account for total taxon diversity. Reliable distribution data and robust phylogenies are needed to further investigate

geographic patterns of diversification (Willmott et al. 2001). Progress is being made with mtDNA and nuclear loci in the ithomiines, at both the genus level and species level, but, it can be argued that it is the more recently diverged taxa that most accurately retain the signatures of the factors that drove the diversifications. Subspecies within ithomiine species are often at least partly reproductively isolated, for example by mimetic allegiance or habitat adaptations, therefore, to test the predictions of distinct mechanisms of diversification we may need to look closely at subspecies level phylogenies.

In summary, mtDNA sequence data revealed that the diversification of 5 subspecies of *H. anchiala* was more consistent with allopatric divergence than ecological adaptation, and occurred largely during the Pleistocene. Evidence suggests that, at least in this group, the divergent forms here assigned to subspecies are evolving most rapidly in the Amazon basin.

References

Barraclough, T. G. & Nee, S. (2001) Phylogenetics and speciation. *Trends in Ecology and Evolution*, **16**: 391-399.

Beccaloni, G. (1997a) Vertical stratification of ithomiine butterfly (Nymphalidae: Ithomiinae) mimicry complexes: the relationship between adult flight height and larval host-plant height. *Biological Journal of the Linnean Society*, **62**: 313-341.

Beccaloni, G. W. (1997b) Ecology, natural history and behaviour of ithomiine butterflies and their mimics in Ecuador (Lepidoptera: Nymphalidae: Ithomiinae). *Tropical Lepidoptera*, **8**: 103-124.

Benson, W. W. (1982) Alternative models for infrageneric diversification in the humid tropics: tests with passion vine butterflies. In *Biological Diversification in the Tropics* (ed. G. T. Prance), pp. 608-640. New York, NY: Columbia Univ. Press.

Brower, A. V. Z. (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**: 6491-6495.

Brower, A. V. Z. (1996) Parallel race formation and the evolution of mimicry in *Heliconius* butterflies: a phylogenetic hypothesis from mitochondrial DNA sequences. *Evolution*, **50**: 195-221.

Brower, A. V. Z., DeSalle, R. & Vogler, A. (1996) Gene trees, species trees, and systematics. A cladistic perspective. *Annual Review of Ecology and Systematics*, **27**: 423-450.

Brown, K. S. (1979) *Ecologia Geográfica e Evolução nas Florestas Neotropicais*. Campinas, Brazil: Universidade Estadual de Campinas.

Brown, K. S. (1982) Historical and ecological factors in the biogeography of aposematic Neotropical butterflies. *American Zoologist*, **22**: 453-471.

Brown, K. S. (1987) Biogeography and evolution of neotropical butterflies. In *Biogeography and Quaternary History in Tropical America* (ed. T. C. Whitmore & G. T. Prance), pp. 66-104. Oxford, U.K.: Oxford Univ. Press.

Brown, K. S. & Benson, W. W. (1974) Adaptive polymorphism associated with multiple Müllerian mimicry in *Heliconius numata* (Lepid.: Nymph.). *Biotropica*, **6**: 205-228.

Endler, J. A. (1982) Pleistocene forest refuges: fact or fancy? In *Biological Diversification in the Tropics* (ed. G. T. Prance), pp. 641-657. New York: Columbia Univ. Press.

Fjeldså, J. (1994) Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodiversity and Conservation*, **3**: 207-226.

Haffer, J. (1969) Speciation in Amazonian forest birds. *Science*, **165**: 131-137.

Hall, J. P. W. (2006) Montane speciation patterns in *Ithomiola* butterflies (Lepidoptera: Riodinidae): are they consistently moving up in the World? Prepared for *Proceedings of the National Academy of Sciences of the United States of America*.

Hall, J. P. W. & Harvey, D. J. (2002) The phylogeography of Amazonia revisited: new evidence from riodinid butterflies. *Evolution*, **56**: 1489-1497.

Huelsenbeck, J. P. & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**: 754-755.

Jiggins, C. D., Naisbit, R. E., Coe, R. L. & Mallet, J. (2001) Reproductive isolation caused by colour pattern mimicry. *Nature (London)*, **411**: 302-305.

Jiggins, C. D., Mallarino, R., Willmott, K. R. & Bermingham, E. (2006) What can phylogenies tell us about speciation? The case of *Ithomia* (Lepidoptera; Nymphalidae). Prepared for *Evolution*.

Joron, M. & Mallet, J. (1998) Diversity in mimicry: paradox or paradigm? *Trends in Ecology and Evolution*, **13**: 461-466.

Joron, M., Wynne, I. R., Lamas, G. & Mallet, J. (1999) Variable selection and the coexistence of multiple mimetic forms of the butterfly *Heliconius numata*. *Evolutionary Ecology*, **13**: 721-754.

Lamas, G. (2004) Nymphalidae. Ithomiinae, pp. 172-191. In G. Lamas (ed.), Checklist: Part 4A. Hesperioidea - Papilionoidea. In J. B. Heppner, (ed.), Atlas of Neotropical Lepidoptera. Volume 5A. Association for Tropical Lepidoptera, Scientific Publishers, Gainesville.

Maddison, W. P. (1997) Gene trees in species trees. *Systematic Biology*, **46**: 523-536.

Müller, F. (1879) *Ituna* and *Thyridia*; a remarkable case of mimicry in butterflies. *Transactions of the Entomological Society of London*, **1879**: xx-xxix.

Posada, N. M. & Crandall, K. A. (1998) *MODELTEST*: testing the model of DNA substitution. *Bioinformatics*, **14**: 817-818.

Robbins, R. K. & Opler, P. A. (1996) Butterfly diversity and a preliminary comparison with bird and mammal diversity. In *Biodiversity II. Understanding and Protecting our Biological Resources* (ed. M. L. Reaka-Kudla, D. E. Wilson & E. O. Wilson), pp. 69-82. Washington, DC: National Academy of Sciences.

Sheppard, P. M., Turner, J. R. G., Brown, K. S., Benson, W. W. & Singer, M. C. (1985) Genetics and the evolution of muellerian mimicry in *Heliconius* butterflies. *Philosophical Transactions of the Royal Society of London Series B- Biological Sciences*, **308**: 433-613.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am., **87**: 651-702.

Swofford, D. L. (2000) *PAUP**: Phylogenetic Analysis Using Parsiomony (*and Other Methods). Sunderland, Massachusetts: Sinauer Associates.

Willmott, K. R. (2001) *The genus Adelpha: its systematics, biology and biogeography.* Gainesville, Florida: Scientific Publishers.

Willmott, K., Hall, J. P. W. & Lamas, G. (2001) Systematics of *Hypanartia* (Lepidoptera: Nymphalidae: Nymphalinae), with a test for geographical speciation mechanisms in the Andes. *Systematic Entomology*, **26**: 369-399.

CHAPTER SIX

A MOLECULAR PHYLOGENY OF THE NEOTROPICAL BUTTERFLY TRIBE OLERIINI (LEPIDOPTERA: NYMPHALIDAE: ITHOMIINAE)

Abstract

Here I present gene genealogies for 41 of the approximately 60 butterfly species of the tribe Oleriini (Lepidoptera: Nymphalidae: Ithomiinae), based on data from three different gene regions; 1072 base pairs (bp) of *elongation factor 1*-α, 417 bp of *wingless*, and 1619 bp of the mitochondrial genes *cytochrome oxidase I* and *II*, plus the intervening tRNA. In all but one case (*Hyposcada taliata* for the mitochondrial region), species within the Oleriini clustered into the appropriate four genera: *Hyposcada*, *Megoleria*, *Ollantaya* and *Oleria*. Neighbour-joining analyses of individual gene regions and combined data, and maximum parsimony analysis of combined data recovered *Hyposcada* as sister to all other Oleriini. However a Bayesian analysis of combined data identified *Megoleria* as sister to all other Oleriini. The latter arrangement is consistent with previous findings based on morphological and molecular evidence. Species within *Oleria* have previously been assigned to groups based on morphological characters. The phylogenetic hypotheses presented here suggest that the 'onega' species group is valid, but that additional molecular data are required to determine validity of the other species groups. Further molecular work, and the incorporation of morphological evidence, will enable a complete and robust species-level phylogeny to be generated for this tribe. This will be coupled with detailed distributional data to investigate patterns of diversification across the Oleriini.

Introduction

This chapter marks the beginning of a long-term, collaborative project on the Oleriini butterfly tribe (Lepidoptera: Nymphalidae: Ithomiinae). Until very recently the systematics of the Oleriini were poorly understood, in part due to taxonomic difficulties of cryptic morphology, undescribed life stages, and differences between the sexes. These difficulties resulted in a number of classifications being proposed, for example, Fox (1968) placed the *Hyposcada*, *Oleria* and *Aeria* in the Oleriini. However, based on larval morphology and host associations, *Aeria* has since been recognised as a member of a distinct tribe (Ackery 1988, Brown & Freitas 1994). Motta (2003) moved *Oleria*, *Ollantaya*, *Megoleria*, and *Hyposcada*, as well as most of the genera recognised as belonging to the Napeogenini and Ithomiini by Lamas (2004), into a larger Ithomiini tribe. For the purpose of this chapter, Oleriini is regarded as a monophyletic tribe containing four genera; *Hyposcada*, *Oleria*, and the more recently described *Ollantaya* and *Megoleria*. This is in accordance with classifications by Harvey (1991), and Lamas (2004), and also with molecular data (Brower et al. 2005).

A major aim of the Oleriini project is to describe the genealogical relationships between all species within the Oleriini. These molecular systematic findings, in conjunction with detailed morphological and ecological information currently being accumulated by K. R. Willmott, will form the basis of a thorough revision of the Oleriini tribe.

When complete and robust species-level phylogenies of the Oleriini have been established, they will be coupled with detailed distributional data to test between hypotheses of Neotropical diversification. A number of features of this tribe make it particularly suitable for such analysis. First, the Oleriini is very speciose, and contains many taxa with restricted geographic ranges. In addition, it was predicted to contain some members resulting from recent as well as older radiations. These factors suggest that members of the Oleriini might have retained important signatures, such as the geographical context, of their speciation. Secondly, the Oleriini are distributed from Mexico to Southern Brazil, and therefore provide an opportunity to test between theories of diversification both on a local scale and from a wider phylogeographical perspective. Thirdly, members of the Oleriini are distributed over a wide

altitudinal range (0-3000 m). A number of authors have debated whether Tropical biodiversity can largely be attributed to diversifications in montane regions (Fjeldsä et al. 1994, Roy et al. 1997) or lowland regions (Brown 1957, Barkman & Simpson 2001). Oleriini diversity peaks in the Ecuadorian and Peruvian Andes, as well as in nearby Amazonian lowlands, so the Oleriini provide a model group for testing the relative contributions made by different elevations to Neotropical diversity.

Here, I use data from three different gene regions, *elongation factor 1-α* (*Ef1α*), *wingless* (*wg*), and the mitochondrial genes *cytochrome oxidase I* and *II* (COI and COII) plus their intervening tRNA, in order to infer species level genealogies representing 41 of the approximately 60 species belonging to the Oleriini.

Materials and methods

Taxonomic sampling and DNA extraction

103 individuals representing 41 species of the Oleriini were examined (7 *Hyposcada*, 2 *Ollantaya*, 2 *Megoleria* and 30 *Oleria*) (Table 1). These samples were collected by the author and preserved in 20% dimethylsulphoxide, 0.25M EDTA and saturated NaCl solution, or donated by collaborators as dried specimens or preserved in alcohol. Typically DNA was extracted from one third of a thorax (but sometimes from just one or two meso- and metathoracic legs) using the DNAeasy kit (QIAgen), according to the manufacturer's instructions, with an initial 3 h incubation at 55°C, and a final elution volume of 300 µl. In addition, DNA from three samples obtained from the Brower laboratory (E32-1, B16-7, E16-3) had previously been amplified. Sequence data for two ithomiine outgroups, *Ithomia drymo* and *Hyalyris antea*, were obtained from Genbank (Brower et al. 2005).

Primer development, PCR and sequencing

PCRs were typically performed in a 50 µl volume, using 4 µl template DNA, 1 x PCR buffer (0.1 M Tris HCl, 0.5 M KCl, 0.01 volumes Triton X-100), 0.2µM dNTPs, 0.3 U/µl Taq and; for COI and COII: 16 µM $MgCl_2$, 0.08 µM each primer; for *wg* and *Ef1α*: 5 µM $MgCl_2$, 0.2 µM each primer (Table 2); with

an amplification profile of; a 2 min denaturation at 94°C, followed by; COI: 32 cycles of (94°C for 60 sec, 60°C for 60 sec, 72°C for 90 sec); COII: 33 cycles of (94°C for 60 sec, 56°C for 60 sec, 72°C for 90 sec); *wg:* 30 cycles of (94°C for 60 sec, 62.5°C for 60 sec, 72°C for 60 sec), increased to 35 cycles with a 57.5°C annealing temperature if quality of template was poor; *Ef1α:* 30 cycles of (94°C for 60 sec, 62°C for 60 sec, 72°C for 90 sec); all then followed by a final 10 min extension at 72°C.

Depending on the quality of the PCR products, they were either purified directly using the QIAquick PCR purification kit (QIAgen), or excised from an agarose gel and purified using the QIAquick gel extraction kit (QIAgen), according to the manufacturer's protocols. Purified products were sent to a commercial company (Macrogen, Korea) or cycle-sequenced in a MJ Research DNA Engine, precipitated using ethanol or ethanol and sodium acetate, and run on an ABI 373 automated sequencer using ABI protocols with the PCR primers. In cases where chromatograms did not read for the full sequence length, sequencing was also performed with additional sequencing primers (Table 2).

Data analyses

Sequences were edited using SeqEd v1.0.3 software (Applied Biosystems). PAUP version 4.0b 10 (Swofford 2000) was used to calculate the numbers of variable and parsimony informative sites, nucleotide composition, and to perform a chi-square test of base frequency homogeneity for each gene region. Due to computing constraints, Neighbour Joining (NJ) analysis only was performed in PAUP for each gene region individually, as well as for combined data. Two combined data sets were analysed; 'total' data which included all available data, and 'complete' data which included only those individuals for which data were available for all three gene regions. NJ trees were generated using HKY85 distances with random break ties. Confidence in each node was assessed by performing 1000 rounds of NJ bootstrapping.

In addition, maximum parsimony (MP) topologies were conducted on the 'complete' data set in PAUP using a 10-replicate heuristic search with tree-bisection-reconnection (TBR) branch swapping, with random stepwise addition of sequences. These were summarized with a 95% majority rule consensus tree.

To dissect the contributions of the individual genes in recovering relationships at different taxonomic levels in the 'complete' data set, partitioned Bremer support values (Baker & DeSalle 1997) were calculated using TreeRot (Sorenson 1999). Bayesian analysis was also performed on the 'complete' data set with just one outgroup taxon (MrBayes 3.0, Huelsenbeck & Ronquist 2001) using parameters identified by Modeltest 3.04 (nst=6, rates=invgamma), with four simultaneous chains run for 1,000,000 generations, sampling a tree every 100 generations. A consensus tree with branch support in the form of posterior probabilities was derived from the final 9,000 trees (representing the final 900,000 generations), after confirmation that likelihood values had stabilized.

Results and Discussion

Molecular results

Sequence data were obtained from 103 individuals of the Oleriini; 67 individuals for *Ef1α*, 67 for *wg*, and 96 for the mitochondrial region. The final alignment spans 1072 bp of *Ef1α*, 417 bp of *wg*, and 1619 bp of mitochondria CO region. As expected, the mitochondrial region evolved the fastest: excluding the outgroup taxa; 156 (14.6%) sites were variable for *Ef1α*, of which 102 were parsimony informative (PI); 94 (22.5%) sites were variable (72 PI) for *wg*; and 576 (35.6%) sites were variable (479 PI) for the mitochondrial region. Of 45 branches with Bremer support values (Figure 1), 35, 4 and 6 were most strongly support by the mitochondrial, *Ef1α* and *wg* partitions respectively (an additional branch was equally supported by all partitions). This indicates that the mitochondrial region provided the strongest support for relationships within the Oleriini, consistent with the slow evolutionary rate of the nuclear regions. The partitioned Bremer analysis also revealed that, in general, the nuclear regions provided stronger support to the deeper nodes, consistent with likely saturation of the faster evolving mitochondrial region. As previously found (Whinnett et al. 2005), nucleotide bias was strongest in the mitochondrial region, which had 33.0: 13.5: 11.7: 41.7% A:C:G:T, compared to 25.4: 26.5: 23.8: 24.2 %, and 25.2: 26.1: 27.4: 21.3% A:C:G:T in *Ef1α* and *wg* respectively. Chi square tests for base frequency homogeneity revealed no differences across taxa (P= 1.00) for the three gene regions.

Table 1. Name and identification numbers of the *Hyposcada* (*H.*), *Megoleria* (*M.*), *Ollantaya* (*Ollantaya*), and *Oleria* (*O.*) specimens sequenced, plus the outgroup specimens *Ithomia* and *Hyalyris*. All known information about: the collection locality and co-ordinates, sample source (name of collector or donator, plus collection date), as well as the availability and location of the wing voucher, tissue and DNA template, is indicated. Available GenBank accession codes are given, with 'sequenced' denoting cases where sequence data have been generated but not yet submitted to GenBank, and – indicating cases where sequence data have not yet been obtained.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
H. anchiala ecuadorina	Ec 452	ECUADOR: Napo: Jatun Sacha. 01°04′S, 77°36′W	DQ078474	DQ143816	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Ecuador collection)	UCL (Oleriini collection)
H. anchiala fallax	02-3519	PERU: Madre de Dios: Mazuko. 13°06′S, 70°22′W	DQ078477	DQ085434	DQ085446	J. Mallet et al., 2002	missing	UCL (Peru collection 2002)	UCL (Oleriini collection)
H. anchiala interrupta	02-512	PERU: San Martín: km 7.2 Pongo-Barranquita. 06°17′20″S, 76°13′41″W	DQ078355	DQ143817	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
H. anchiala interrupta	02-2105	PERU: San Martín: km 7.2 Pongo-Barranquita. 06°17′20″S, 76°13′41″W	DQ078356	DQ085433	DQ085445	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
H. anchiala interrupta	02-1293	PERU: San Martín: km 7.2 Pongo-Barranquita. 06°17′20″S, 76°13′41″W	DQ078312	DQ143818	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
H. anchiala mendax	02-1644	PERU: San Martín: Puente Serranoyacu. 05°40′S, 77°40′W	DQ078357	DQ085435	DQ085447	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
H. anchiala mendax	02-1645	PERU: San Martín: Puente Serranoyacu. 05°40′S, 77°40′W	DQ078358	DQ143819	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
H. anchiala mendax	02-716	PERU: San Martín: Puente Serranoyacu. 05°40′S, 77°40′W	DQ078359	DQ143820	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
H. anchiala mendax	02-1602	PERU: San Martín: La Antena. 06°27′18″S, 76°17′54″W	DQ078360	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
H. anchiala mendax	02-2141	PERU: San Martín: La Antena. 06°27′18″S, 76°17′54″W	DQ078361	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
H. anchiala richardsi	G1 (cp 8-17)	PERU: Madre de Dios: Río Los Amigos. 12°35′S, 70°5′W	DQ078475	-	-	C. Peña & G. Lamas	MHN, Lima	UCL (Oleriini collection)	UCL (Oleriini collection)
H. illinissa napoensis	K6 (KRW-11/2/04-6)	ECUADOR: Napo, Río Yuturi, Río Manduro trail, Río Napo. 00°33.49'S,76°2.65'W	sequenced	-	-	K. R. Willmott, 2004	K. R. Willmott	Extracted whole donated sample	UCL (dried STRI collection)
H. illinissa ida	K7 (KRW-11/2/04-6)	ECUADOR: Napo, Río Yuturi, lodge trail, Río Napo. 00°32.88'S, 76°2.48'W	sequenced	-	-	K. R. Willmott, 2004	K. R. Willmott	Extracted whole donated sample	UCL (dried STRI collection)
H. illinissa ssp nov 4	02-1187	PERU: Km 28, Tarapoto - Yurimaguas, San Martin 06°41′19″S, 76°31′76″W	sequenced	DQ143822	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	None
H. illinissa margarita	02-190	PERU: San Martín: Chumía. 06°36′57″S, 76°11′10″W	sequenced	DQ143823	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	None
H. illinissa	Ec 455	ECUADOR: Napo: Jatun Sacha. 01°04′S, 77°36′W	sequenced	DQ143821	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Ecuador collection)	UCL (Oleriini collection)
H. kena	Ec 458	ECUADOR: Napo: Jatun Sacha. 01°04′S, 77°36′W	sequenced	DQ143825	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Ecuador collection)	UCL (Oleriini collection)
H. kena flexibilis	02-197	PERU: San Martín, Chumía. 06°36′57″S, 76°11′10″W	DQ078362	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
H. kena flexibilis	02-198	PERU: San Martín, Chumía. 06°36′57″S, 76°11′10″W	DQ078363	DQ085436	DQ085448	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
H. kena flexibilis	02-1591	PERU: San Martin, 5 km after Shapaja	DQ078364	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
H. kena ssp nov	02-1666	PERU: San Martín: Puente Serranoyacu. 05°40′S, 77°40′W	DQ078365	DQ143824	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
H. kena ssp nov	02-1670	PERU: San Martín: Puente Serranoyacu. 05°40′S, 77°40′W	DQ078366	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
H. kena ssp nov	02-1667	PERU: San Martín: Puente Serranoyacu. 05°40′S, 77°40′W	DQ078367	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
H. schausi lactea	K5	ECUADOR: Esmeraldas, Río Chuchuví, km 12.5 Lita-San Lorenzo Rd. 00°52.85′N, 78°30.90′W	sequenced	-	-	K. R. Willmott	UCL	Extracted whole donated sample	UCL (dried STRI collection)
H. taliata	02-1696	PERU: San Martin, Rio Nieva. 05°41′17″S, 77°46′54″W	sequenced	DQ143826	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
H. virginiana	8283	PANAMA: Cerro Campana. 08°68′74″N, 79°91′97″W	DQ078476	DQ143828	sequenced	C. Jiggins	Edinburgh	UCL (Oleriini collection)	UCL (Oleriini collection)
H. virginiana evanides	8339	PANAMA: Quebrada Hornito Trail, Fortuna. 08°69′28″N, 82°22′45″W	sequenced	DQ143829	sequenced	C. Jiggins	Edinburgh	Contact C. Jiggins	UCL (Oleriini collection)
H. virginiana	8000	PANAMA: Quebrada Hornito Trail, Fortuna. 08°69′28″N, 82°22′45″W	sequenced	DQ143827	sequenced	C. Jiggins	Edinburgh	UCL (Oleriini collection)	UCL (Oleriini collection)
H. zarepha zarepha	02 K	FRENCH GUIANA: Camp Patawa, route de Kaw	sequenced	DQ143830	-	Stephane Attal	UCL	UCL (Oleriini collection)	UCL (Oleriini collection)

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
M. orestilla	18-IK	18-I ECUADOR: Prov. Zamora-Chinchipe, km 40 Loja- Zamora.	sequenced	DQ143832	sequenced	K. R. Willmott	18-I UCL	18-I Tissue UCL (Oleriini collection)	18-I UCL (Oleriini collection)
M. orestilla	Ec 402	ECUADOR: Yanayacu, near Cosanga	sequenced	DQ143831	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Oleriini collection)	None
M. susiana	M27	ECUADOR: Napo, Cordillera Guacamayos, Baeza-Tena Rd., Cocodrilo area.	-	DQ143833	-	Ismael Aldas	NHM, London	Extracted whole donated sample	None
Ollantaya aeginata	E 32-1	ECUADOR: Prov. Zamora-Chinchipe, Zamora	DQ157526	DQ175515	DQ177972	K. R. Willmott	OSU	OSU	OSU
Ollantaya olerioides	13-19K	ECUADOR: Sucumbíos, Río Sucio, nr. La Bonita. 00°28.50'N, 77°33.30'W,	-	DQ143834	sequenced	K. R. Willmott	UCL	UCL (Oleriini collection)	None
O. agarista agarista	02-294	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078368	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. agarista agarista	02-1307	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078369	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. agarista agarista	02-293	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078370	DQ143836	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
O. alexina didymaea	02-3370	PERU	sequenced	DQ143837	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	None
O. amalda modesta	Ec 135	ECUADOR: near Hotel Tinalandia, Alluriquin.	sequenced	DQ143838	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Oleriini collection)	None
O. aquata	B16-7	BRAZIL: São Paulo, Jundia.	DQ168618	DQ168620	DQ168616	A. Freitas	OSU	OSU	OSU
O. assimilis pilcopata	02-3609	PERU: Cuzco, Quincemil.	DQ085456	DQ085438	DQ085450	J. Mallet et al., 2002	missing	UCL (Peru collection 2002)	missing
O. assimilis	16-8K	ECUADOR: Sucumbíos, El Recodo, La Bonita-Lumbaqui Rd.	sequenced	DQ143839	sequenced	K. R. Willmott	UCL	UCL (Oleriini collection)	UCL (Oleriini collection)
O. athalina banjana	Ec 415	ECUADOR: Yanayacu, near Cosanga.	sequenced	-	-	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Oleriini collection)	None
O. athalina banjana	11-6K	ECUADOR: Sucumbíos, Old Rd. to La Alegría, km 19 La Bonita-Tulcan Rd. 00°33.70'N, 77°31.70'W	-	DQ143840	sequenced	K. R. Willmott	UCL	UCL (Oleriini collection)	UCL (Oleriini collection)
O. attalia	35-1K	ECUADOR: Zamora-Chinchipe, San Francisco, canal subterraneo, Zamora-Loja Rd. 03°58.74'S, 79°5.13'W	sequenced	DQ143841	sequenced	K. R. Willmott	Contact K. R. Willmott	Extracted whole donated sample	UCL (Oleriini collection)
O. baizana	Ec 413	ECUADOR: Yanayacu, near Cosanga.	-	DQ143842	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	None	UCL (Oleriini collection)

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
O. bioculata tapio	02-1611	PERU: San Martín: La Antena. 06°27'18"S, 76°17'54"W	sequenced	DQ143843	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	None
O. boyeri	Boy	VENEZUELA: Estado Bolívar, La Escalera km 129	sequenced	-	sequenced	J. Mavarez	UCL	STRI	STRI
O. cyrene solida	Ec 394	ECUADOR: Yanayacu, near Cosanga	sequenced	DQ143844	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Oleriini collection)/ UCL (Ecuador collection)	UCL (Oleriini collection)
O. estella	02-406	PERU: San Martín: La Antena. 06°27'18"S, 76°17'54"W	sequenced	DQ143845	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	None
O. fasciata fasciata	9-2K	ECUADOR: Sucumbíos, Río Sucio, nr. La Bonita. 00°28.50'N, 77°33.30'W	sequenced	DQ143846	sequenced	K. R. Willmott	UCL	UCL (Oleriini collection)	UCL (Oleriini collection)
O. fumata	K14	W. COLOMBIA	sequenced	-	-	L. Constantino	UCL	UCL (Oleriini collection)	UCL (dried STRI collection)
O. fumata	K13	W. COLOMBIA	sequenced	-	-	L. Constantino	UCL	UCL (Oleriini collection)	UCL (dried STRI collection)
O. gunilla lota	02-1304	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078374	DQ143848	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	None
O. gunilla lota	02-2118	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078375	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. gunilla lota	02-290	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078376	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
O. gunilla serdolis	02-1580	PERU: San Martin, 5 km after Shapaja	DQ078377	DQ143847	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. gunilla serdolis	02-196	PERU: San Martín: Chumía. 06°36'57"S, 76°11'10"W	DQ078379	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. gunilla serdolis	02-194	PERU: San Martín: Chumía. 06°36'57"S, 76°11'10"W	DQ078378	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. gunilla lota	Ec 467	ECUADOR: Napo: Jatun Sacha. 01°04'S, 77°36'W	DQ085457	DQ085439	DQ085451	A. Whinnett & R. Naisbit., 2002	UCL	UCL (Oleriini collection)	None
O. ilerdina lerida	02-1301	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078371	DQ143849	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	UCL (DNA extracts 2005)
O. ilerdina lerida	02-1302	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078372	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. ilerdina lerida	02-2023	PERU: San Martin Km 26, Yurimaguas- Tarapoto	sequenced	DQ143850	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	None
O. ilerdina lerida	02-291	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078373	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. makrena	6-3	ECUADOR: Sucumbíos, km 8 La Bonita-Tulcán Rd. 00°30.90'N, 77°31.50'W,	sequenced	DQ143835	-	K. R. Willmott	Contact K. R. Willmott	UCL (Oleriini collection)	None
O. makrena	K16	W. COLOMBIA	sequenced	-	-	L. Constantino	UCL	UCL (Oleriini collection)	UCL (dried STRI collection)
O. makrena	K15	COLOMBIA: Antioquia, Amagá, Finca el Socorro, Vereda la Florida	sequenced	-	-	L. Gomez Piñerez	UCL	None	UCL (dried STRI collection)

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
O. makrena makrenita	Ec 409	ECUADOR: Yanayacu, near Cosanga	sequenced	DQ143852	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	None	None
O. onega ssp nov	02-1708	PERU: San Martín: Puente Serranoyacu. 05°40′S, 77°40′W	DQ078380	DQ143855	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
O. onega ssp nov	02-1709	PERU: San Martín: Puente Serranoyacu. 05°40′S, 77°40′W	DQ078381	DQ143856	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
O. onega ssp nov	02-1674	PERU: San Martín: Puente Serranoyacu. 05°40′S, 77°40′W	DQ078382	DQ143857	sequenced	J. Mallet et al., 2002	missing	UCL (Peru collection 2002)	UCL (Oleriini collection)
O. onega ssp nov	02-1173	PERU: San Martin, 5 km after Shapaja	DQ078383	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
O. onega ssp nov	02-1585	PERU: San Martin, 5 km after Shapaja	DQ078384	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	(Oleriini collection)
O. onega ssp nov	02-1958	PERU: San Martin, 5 km after Shapaja	DQ078385	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
O. onega janarilla	02-520	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17′20″S, 76°13′41″W	DQ078391	DQ143854	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
O. onega janarilla	02-1165	PERU: San Martin, 5 km after Shapaja	DQ078386	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. onega janarilla	02-521	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17′20″S, 76°13′41″W	X	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	None
O. onega janarilla	02-1583	PERU: San Martin, 5 km after Shapaja	DQ078387	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
O. onega janarilla	02-1960	PERU: San Martin, 5 km after Shapaja	DQ078388	-	-	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (DNA extracts 2005)
O. onega janarilla	02-303	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078389	DQ143853	sequenced	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
O. onega janarilla	02-515	PERU: San Martin, Km 7.2, Pongo- Barranquita. 06°17'20"S, 76°13'41"W	DQ078390	DQ085441	DQ085453	J. Mallet et al., 2002	UCL	UCL (Peru collection 2002)	UCL (Oleriini collection)
O. onega janarilla	Ec 277	ECUADOR: Pastaza, Communidad Shuar Mirador, 70km E of Macas-Puyo	DQ085458	DQ085440	DQ085452	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Oleriini collection)	None
O. onega ssp nov	02-835	PERU: San Martín: Puente Serranoyacu. 05°40'S, 77°40'W	DQ085459	DQ085442	DQ085454	J. Mallet et al., 2002	missing	To check	UCL (Oleriini collection)
O. padilla cajamarcensis	Ec 167	ECUADOR: Vilcabamba Ecolodge, Eastern Trail from Ruinas Lodge to Rumi Wilco	sequenced	DQ143858	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Oleriini collection)	None
O. padilla makrena	02-1699	PERU: San Martin, Rio Nieva. 05°41'17"S, 77°46'54"W	sequenced	DQ143860	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	None
O. padilla padilla	Ec 85	ECUADOR: Mindo	sequenced	DQ143859	sequenced	A. Whinnett & R. Naisbit, 2002	UCL	UCL (Oleriini collection)	None
O. padilla	19-1	ECUADOR: Pichincha, Palmito Pamba, km 5 Nanegal-Garcia Moreno Rd.	-	-	sequenced	K. R. Willmott	UCL	None	UCL (Oleriini collection)
O. paula	8474	PANAMA: Cerro Campana. 08°68'74"N, 79°91'97"W	sequenced	DQ143861	sequenced	C. Jiggins	Edinburgh	Contact C. Jiggins	UCL (Oleriini collection)

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
O. quadrata	E16-3	ECUADOR: Pichincha Prov., San Antonio	DQ168619	DQ168621	DQ168617	K. R. Willmott	OSU	OSU	OSU
O. radina ssp nov	11-2	ECUADOR: Sucumbíos, Old Rd. to La Alegría, km 19 La Bonita-Tulcan Rd. 00°33.70'N, 77°31.70'W	sequenced	DQ143862	sequenced	K. R. Willmott	Contact K. R. Willmott	UCL (Oleriini collection)	None
O. rubescens	8369	PANAMA: Chiriqui, Quebrada Hornito. 08°69'28"N, 82°22'45"W	DQ085460	DQ085443	DQ085455	C. Jiggins	Edinburgh	UCL (Oleriini collection)	UCL (Oleriini collection)
O. rubescens	8404	PANAMA: Continental Divide Trail, Fortuna. 08°78'53"N, 82°21'43"W	sequenced	DQ143863	sequenced	C. Jiggins	Edinburgh	Contact C. Jiggins	UCL (Oleriini collection)
O. santineza	3-2	ECUADOR: Sucumbíos, Río Sucio, nr. La Bonita. 00°28.50'N, 77°33.30'W	DQ157525	DQ175514	DQ177971	K. R. Willmott	Contact K. R. Willmott	UCL (Oleriini collection)	None
O. sexmaculata	M29	ECUADOR: Napo, Estación Científica Yasuní, Río Tiputini, via Auca. 00°38.0'S, 76°25.0'W,	sequenced	-	-	K. R. Willmott & J. Hall	NHM, London	Extracted whole donated sample	None
O. tigilla tigilla	K12	ECUADOR: Napo, Pimpilala, Río Jatunyacu	sequenced	-	-	K. R. Willmott	UCL	Extracted whole donated sample	UCL (dried STRI collection)
O. tremona	6-4K	ECUADOR: Sucumbíos, km 8 La Bonita-Tulcán Rd. 00°30.90'N, 77°31.50'W	sequenced	DQ143851	sequenced	K. R. Willmott	UCL	UCL (Oleriini collection)	UCL (Oleriini collection)
O. vicina	2987	PANAMA: Continental Divide Trail, Fortuna. 08°78'53"N, 82°21'43"W	sequenced	-	-	C. Jiggins	Edinburgh	UCL (Oleriini collection)	None

Table 1; continued.

Taxa	Specimen number	Collection locality	mito-chondrial	*wg*	*Ef1α*	Source	Wing voucher	Available tissue	Available DNA template
O. vicina	8500	PANAMA: Palo Alto, Boquete.	sequenced	DQ143864	sequenced	C. Jiggins	Edinburgh	UCL (Oleriini collection)	UCL (Oleriini collection)
O. vicina	8498	PANAMA: Palo Alto, Boquete.	-	DQ143865	sequenced	C. Jiggins	Edinburgh	UCL (Oleriini collection)/ Contact C. Jiggins	UCL (Oleriini collection)
O. victorine	02-3339	PERU	sequenced	DQ143867	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	None
O. victorine sarilis	02-1458	PERU: San Martín: Chumía. 06°36′57″S, 76°11′10″W	sequenced	DQ143866	sequenced	J. Mallet et al., 2002	UCL	UCL (Oleriini collection)	None
O. zelica pagasa	8396	PANAMA: Continental Divide Trail, Fortuna. 08°78′53″N, 82°21′43″W	sequenced	DQ143868	sequenced	C. Jiggins	Edinburgh	UCL (Oleriini collection)	UCL (Oleriini collection)
sO. zelica	8-1K	ECUADOR: Imbabura, Rumiñahui, 37 km N Pedro Vicente Maldonado. 00°16.73′N, 78°59.9′W	-	-	sequenced	K. R. Willmott	UCL	None	UCL (Oleriini collection)
Ithomia drymo	B16-5	BRAZIL: São Paulo, Jundia	DQ069238	DQ073014	DQ073030	A. Freitas	OSU	OSU	OSU
Hyalyris antea	E30-4	ECUADOR: Prov. Zamora-Chinchipe, Quebrada Chorillos	DQ069237	DQ073013	DQ073029	K. R. Willmott	OSU	OSU	OSU

Table 2. Primers used to amplify mtDNA, *wg* and *Ef1α* fragments in specimens of the Oleriini, indicating; primer name, 5'-3' sequence, gene region, its utility for PCR and sequencing, and the source of the primer. The primers with the source reference 'Whinnett' were designed from Oleriini sequences as Oleriini specific primers, except for Ithom-EF-fo which was designed from a wider range of ithomiines. The primers with the source references: Whinnett (modified from primers in Cho et al. 1995); and Murray (modified from primers in Cho et al. 1995); were designed for general use in Lepidoptera, from primers listed by Cho and colleagues (1995). Note that the primers with source references Whinnett, Whinnett (modified) or Murray (modified) have not yet been published elsewhere.

Primer (name as ordered in OSU)	Region	5'-3' Sequence	Source	PCR primer	Sequencing primer
Jerry-COI	COI	CAACATTTATTTTGATTTTTTGG	Simon et al. 1994	Y	Y
PatCOI/ Pat2COI	COI	TCCATTACATATAATCTGCCATATTAG	Simon et al. 1994	Y	Y
Pat-aw	COI	TCCAATGCACTAATCTGCCATATTA	Simon et al. 1994	Y	Y
OleriiniCOInest	COI	TGATATCATTCAATAGAWGAWGGT	Whinnett		Y
COI-int-Oler-aw	COI	ACTCCWGTTAATCCTCCNACA	Whinnett		Y
George3-aw	COII	TAGGTITAGCIGGAATACCTCG	Simon et al. 1994	Y	Y
George IV	COII	TTTAGCTGGAATACCTCGACG	Simon et al. 1994	Y	Y
Imelda/Imelda-aw	COII	CATTAGAAGTAATTGCTAATTTACTA	Simon et al. 1994	Y	Y
COII-Ol-intfo-aw	COII	CANGARTGAATAACATCAGTAGCAG	Whinnett		Y
COII-Ol-intrev-aw-r	COII	CTGCTACTGATGTTATTCAYTCNTG	Whinnett		Y
WG1	*wg*	GARTGYAARTGYCAYGGYATGTCTGG	Brower and DeSalle 1998	Y	Y
WG2	*wg*	ACTICGCRCACCARTGGAATGTRCA	Brower and DeSalle 1998	Y	Y
EF1a-fo-aw	*Ef1α*	CTGAGCGYGARCGTGGTAT	Whinnett (modified from primers in Cho et al. 1995)	Y	
EF1a-rev-aw	*Ef1α*	ACAGCNACKGTYTGYCTCAT	Whinnett (modified from primers in Cho et al. 1995)	Y	Y
Ithom-EF-fo	*Ef1α*	TCGAAACCGCTAARTTCTATGT	Whinnett		Y
EF-int-Oleriini-aw	*Ef1α*	ATCGCTCTGTGGAAGTTCG	Whinnett		Y
EF-Ol-fo2-aw	*Ef1α*	AGACAACGTAGGTTTCAACG	Whinnett		Y
EF-1alphaM46Al	*Ef1α*	GAGGAAATYAARAAGGAAG	Murray (modified from primers in Cho et al. 1995)		Y
EF-1alphaM51Monica	*Ef1α*	CATRTTGTCKCCGTGCCAKCC	Murray (modified from primers in Cho et al. 1995)		Y

Generic level phylogenetics

With the exception of *H. taliata* for the mitochondrial region, NJ analysis revealed all individuals to cluster by species, for the individual gene and 'complete' data sets. For genera, the phylogenetic hypotheses inferred for the nuclear regions revealed the *Hyposcada* as sister to a group containing *Oleria, Megoleria* and *Ollantaya*. The *wg* data revealed *Megoleria* as sister to *Ollantaya* + *Oleria* (Figure 2), whilst the *Ef1 α* data recovered *Oleria* as sister to *Megoleria* + *Ollantaya* pairing (Figure 3). The hypothesis inferred using mitochondrial DNA recovered *Hyposcada* (except *H. taliata*) as sister to all other Oleriini individuals. The mitochondrial region further revealed *Ollantaya* as sister to *Oleria* + (*Megoleria* + *H. taliata*) (Figure 4). Given these topological conflicts, I investigated which relationships emerged when the data were combined in single analyses. The 'complete' phylogenetic hypothesis most closely mirrored *wg* at the generic level, revealing the genus *Hyposcada* as sister to a (*Megoleria* (*Ollantaya* + *Oleria*)) grouping (Figure 5).

The repeated finding here, that *Hyposcada* is sister to all other Oleriini, contradicts previous reports. Based on morphology, Brown and Freitas (1994), reported *Ollantaya* as the sister genus to all other Oleriini. In contrast, but also based on morphology, Willmott (personal correspondence) argued that *Megoleria* are sister to all other Oleriini.

Interestingly, Brower and colleagues (2005) also recovered *Megoleria* as the sister species to all other Oleriini, in a higher level phylogeny of the Ithomiinae based on the three gene regions studied here, sampling fewer species from each genus. It is likely that sampling of species, and perhaps different phylogenetic reconstruction methods, account for the different generic arrangements recovered in the two studies.

Although there were many cases of agreement between the phylogenetic hypothesis inferred for the 'total' data set (Figure 6) and hypotheses for the individual gene and 'complete' data sets, there were also some striking differences. For example, the *Megoleria* and *Ollantaya* were paraphyletic with respect to individuals belonging to five species of *Oleria* (*O. athalina, O. baizana, O. padilla, O. zelica* and *O. vicina*) using the 'total' data set . Taking the above mentioned clade, although both *M. orestilla*, the *O. aeginata* and one of the

O. vicina individuals are represented by all three gene regions the other seven individuals are missing mitochondrial sequence data, and three of these are only represented by data for one nuclear gene region. To focus on *O. zelica*, individuals 8396 and 8-IK assume very different positions in the 'total' data set, even though they are identical over the 1072 bp of *Ef1 α*. 8396 assumes a more expected position within the larger *Oleria* clade close to individuals of *O. rubescens*, *O. paula* and *O. amalda*, than the 8-1K which is found within the *Megoleria-Ollantaya* grouping. These different positions are almost certainly due to 8396 having the full complement of gene regions, whilst 8-1K is only represented by *Ef1 α* and is therefore likely confounded by the missing mitochondrial and *wg* data. As the 'total' data set includes a considerable amount of missing data (36 missing *Ef1 α*, 36 missing *wg* and 7 missing mitochondrial sequences) confidence can't be assigned to the phylogeny derived from this alignment, so following discussions will concentrate on the individual gene and 'complete' data sets. More molecular work on the Oleriini is planned for the future, and I strongly recommend sequencing each individual for all 3 gene regions, to improve the robustness of the phylogenies bases on 'total' combined data.

Species level phylogenetics

There are cases of agreement between phylogenies of species within *Hyposcada*, for example; *Ef1 α*, *wg*, mitochondrial and 'complete' data sets all recover *H. anchiala*, *H. kena* and *H. virginiana* as a grouping; in addition, *wg*, *Ef1 α* and 'complete' data sets similarly recover a (*H. illinissa* (*H. taliata* (*H. anchiala* (*H. kena*, *H. virginiana*)))) arrangement. There are also a number of phylogenetic discordances. For example, mitochondrial DNA suggests *H. kena* is closer to *H. anchiala* than to *H. virginiana*; the latter two are grouped by the nuclear genes. In addition, *H. taliata* is clustered closer to *Megoleria* than other *Hyposcada* using mitochondrial DNA, making *Hyposcada* paraphyletic with respect to the other Oleriini. Interestingly, *H. taliata* has the longest basal node to tip branch length on the mitochondrial tree, thus an erroneous position of *H. taliata* in some of these trees might be attributable to long branch attraction.

Figure 1. 'Complete' data phylogenetic hypothesis of 60 Oleriini specimens, representing the 4 genera (*Hyposcada*, *Megoleria*, *Ollantaya* and *Oleria*), plus 2 outgroup individuals (*Ithomia* and *Hyalyris*). Fragment lengths are as follows: 417 bp *wg*, 1072 bp *Ef1α* and 1618 bp mtDNA. Only those individuals for which data were available for all 3 gene regions are included. Tree topology was inferred with maximum parsimony. Partitioned Bremer support values are given above branches (*Ef1α*/ *wg*/ mitochondrial).

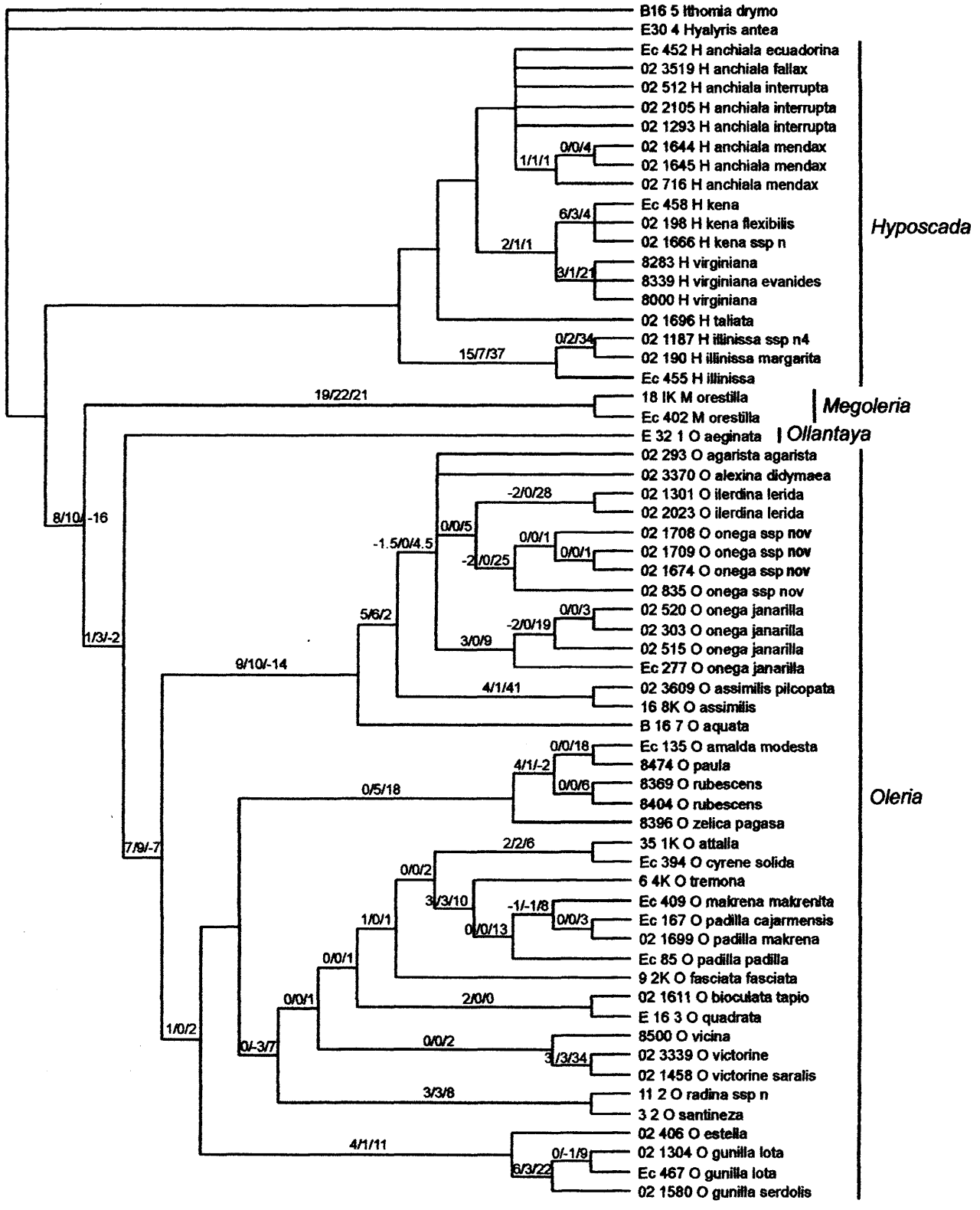


Figure 2. Phylogenetic hypothesis based on 417 bp *wg* data for 67 Oleriini, representing the 4 genera (*Hyposcada*, *Megoleria*, *Ollantaya* and *Oleria*), plus 2 outgroup individuals (*Ithomia* and *Hyalyris*). Tree topology was inferred with neighbour joining. Neighbour joining bootstrap values are given along branches.

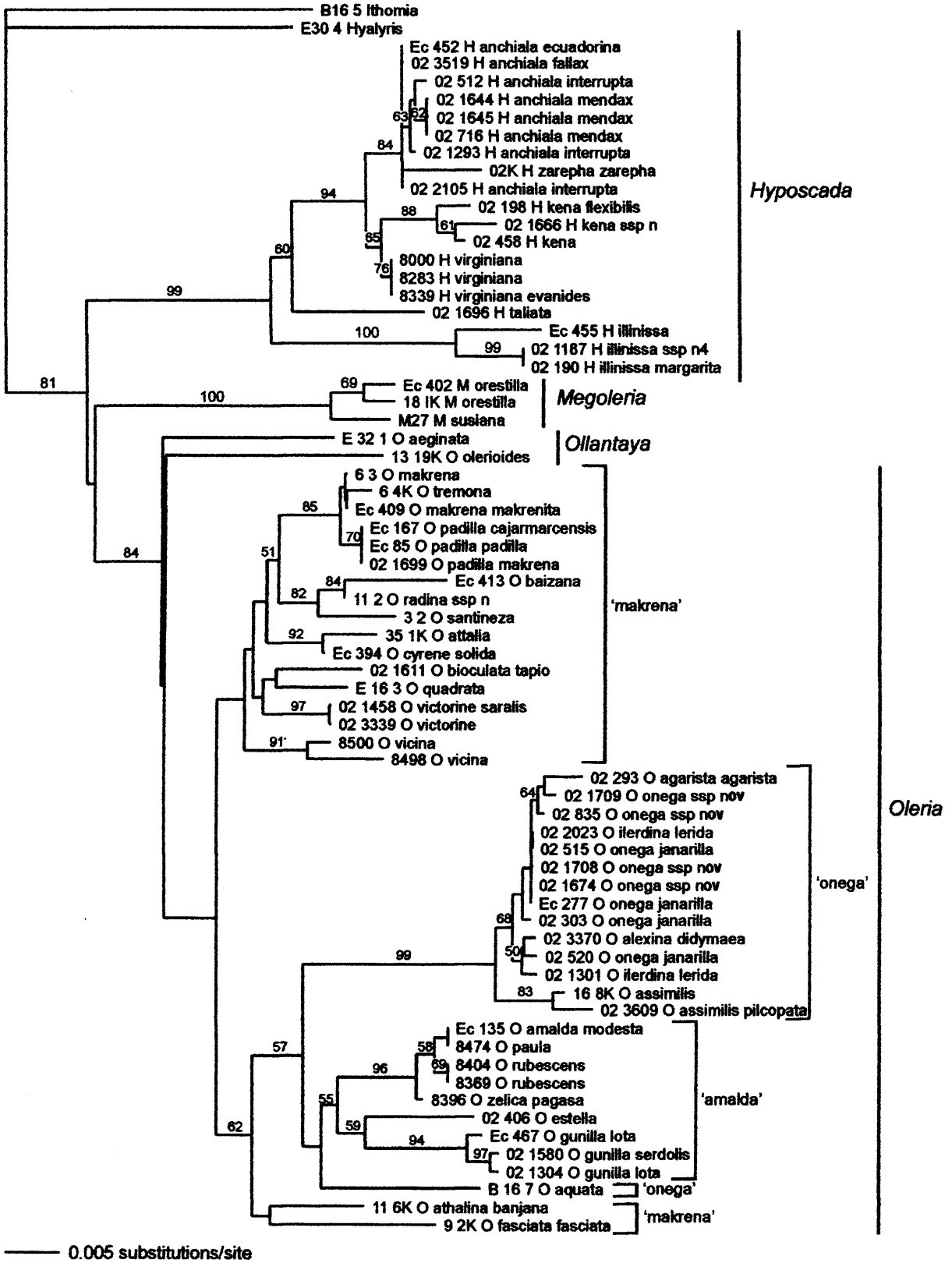


Figure 3. Phylogenetic hypothesis based on 1072 bp *Ef1α* data for 67 Oleriini, representing the 4 genera (*Hyposcada*, *Megoleria*, *Ollantaya* and *Oleria*), plus 2 outgroup individuals (*Ithomia* and *Hyalyris*). Tree topology was inferred with neighbour joining. Neighbour joining bootstrap values are given along branches.

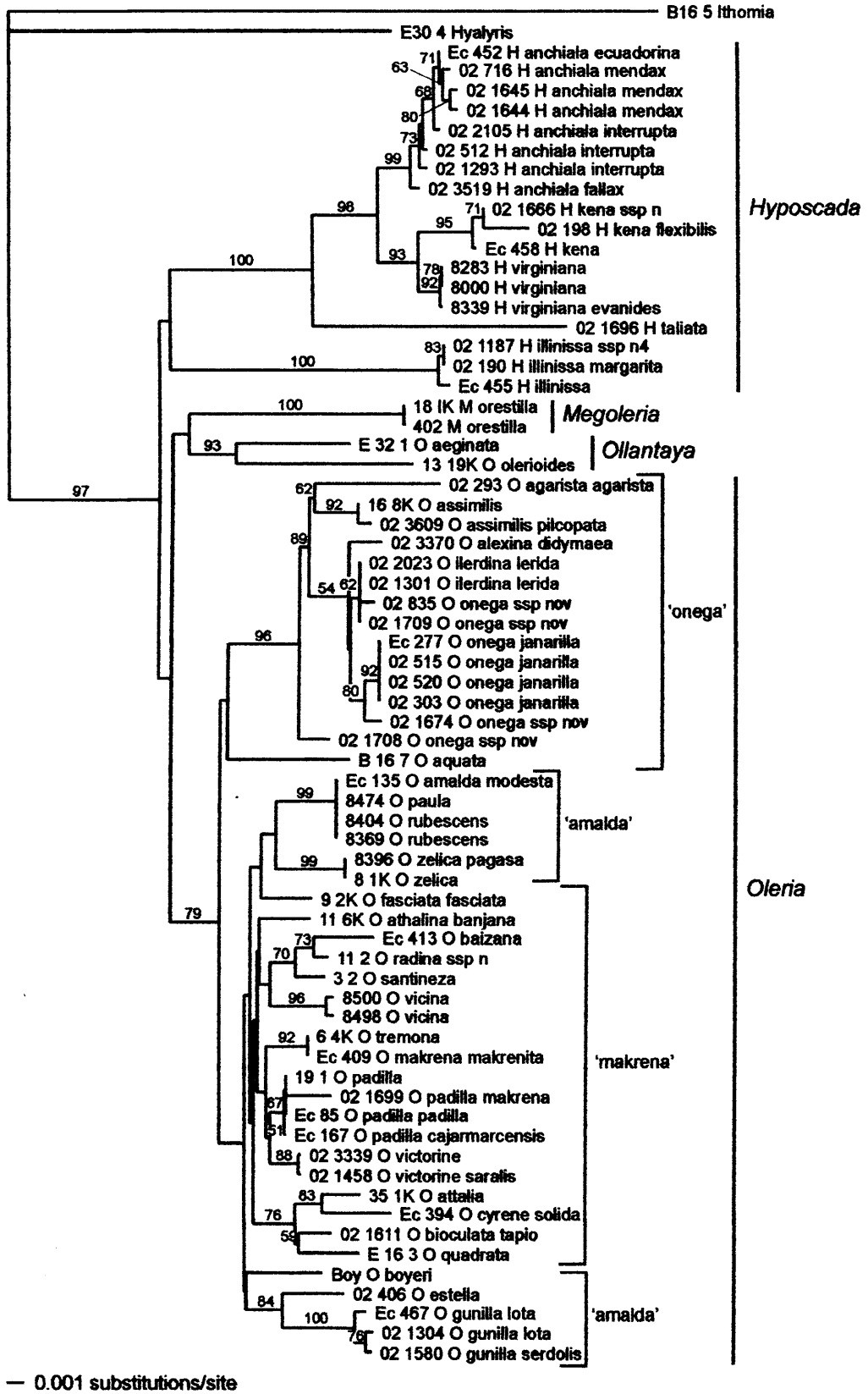


Figure 4. Phylogenetic hypothesis based on 1619 bp mitochondrial data for 96 Oleriini, representing the 4 genera (*Hyposcada*, *Megoleria*, *Ollantaya* and *Oleria*), plus 2 outgroup individuals (*Ithomia* and *Hyalyris*). Tree topology was inferred with neighbour joining. Neighbour joining bootstrap values are given along branches.

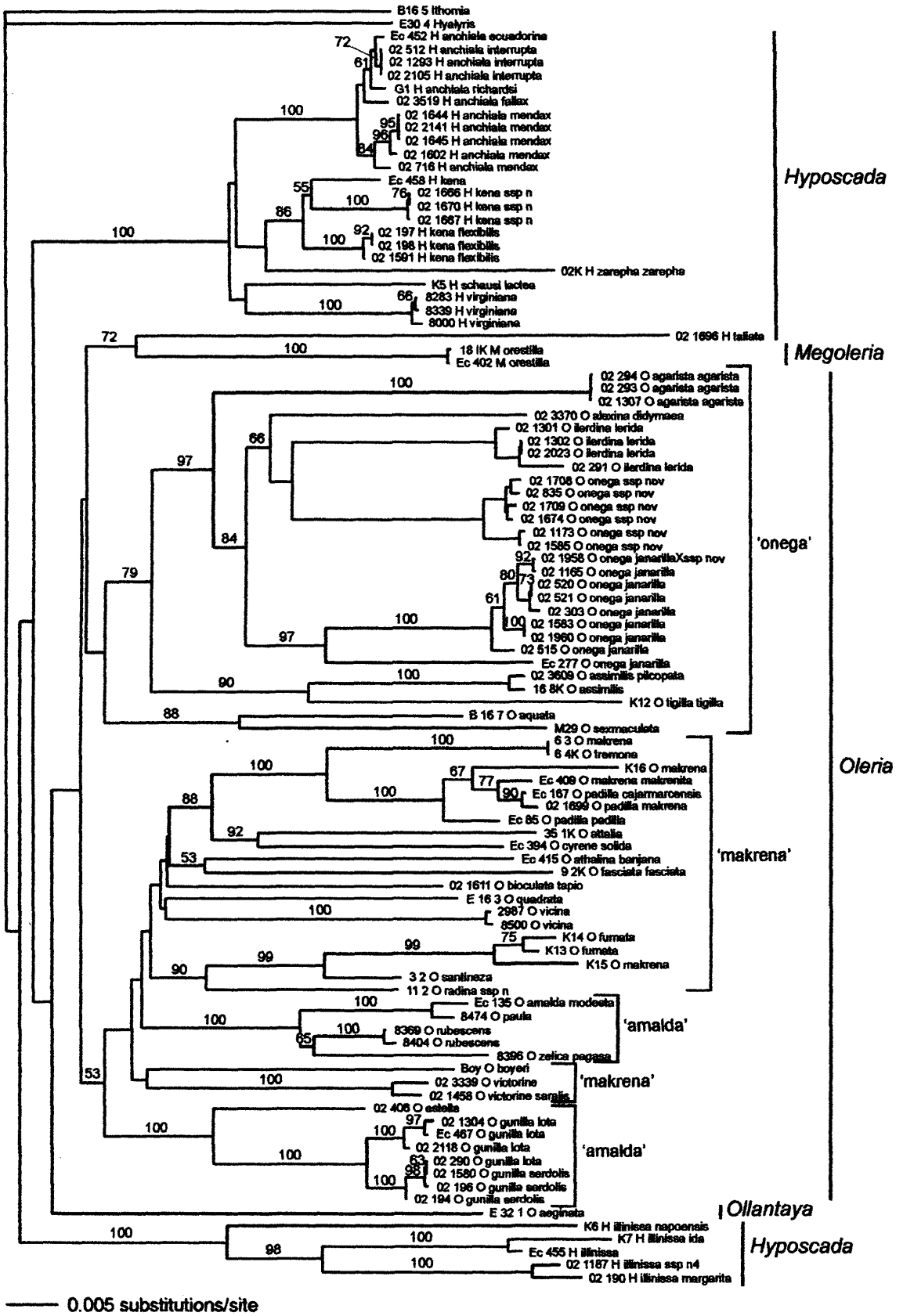


Figure 5. 'Complete' data phylogenetic hypothesis of 60 Oleriini specimens, representing the 4 genera (*Hyposcada*, *Megoleria*, *Ollantaya* and *Oleria*), plus 2 outgroup individuals (*Ithomia* and *Hyalyris*). Fragment lengths are as follows: 417 bp *wg*, 1072 bp *Ef1α* and 1618 bp mtDNA. Only those individuals for which data were available for all 3 gene regions are included. Tree topology was inferred with neighbour joining. Neighbour joining bootstrap values are given along branches.

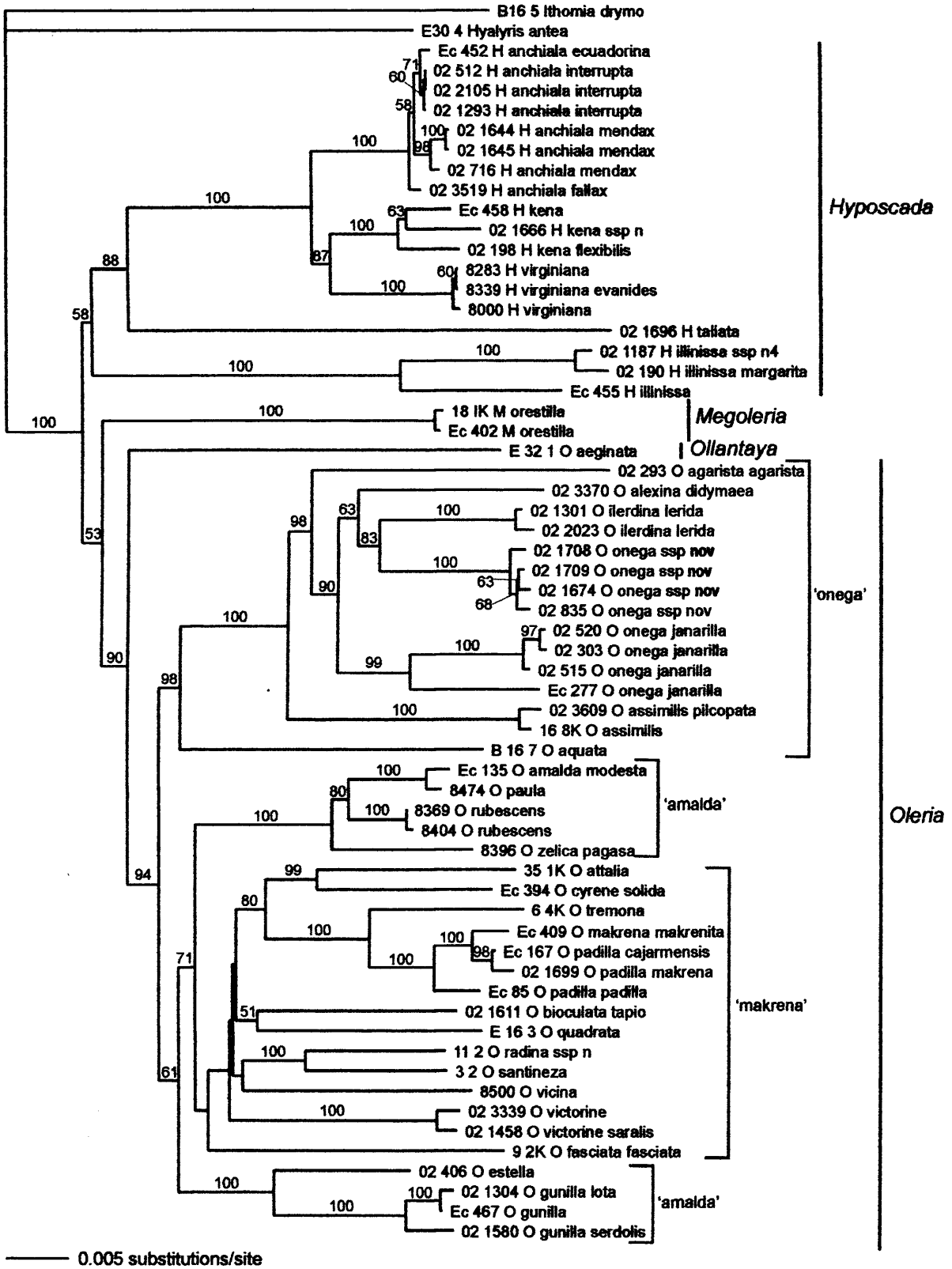
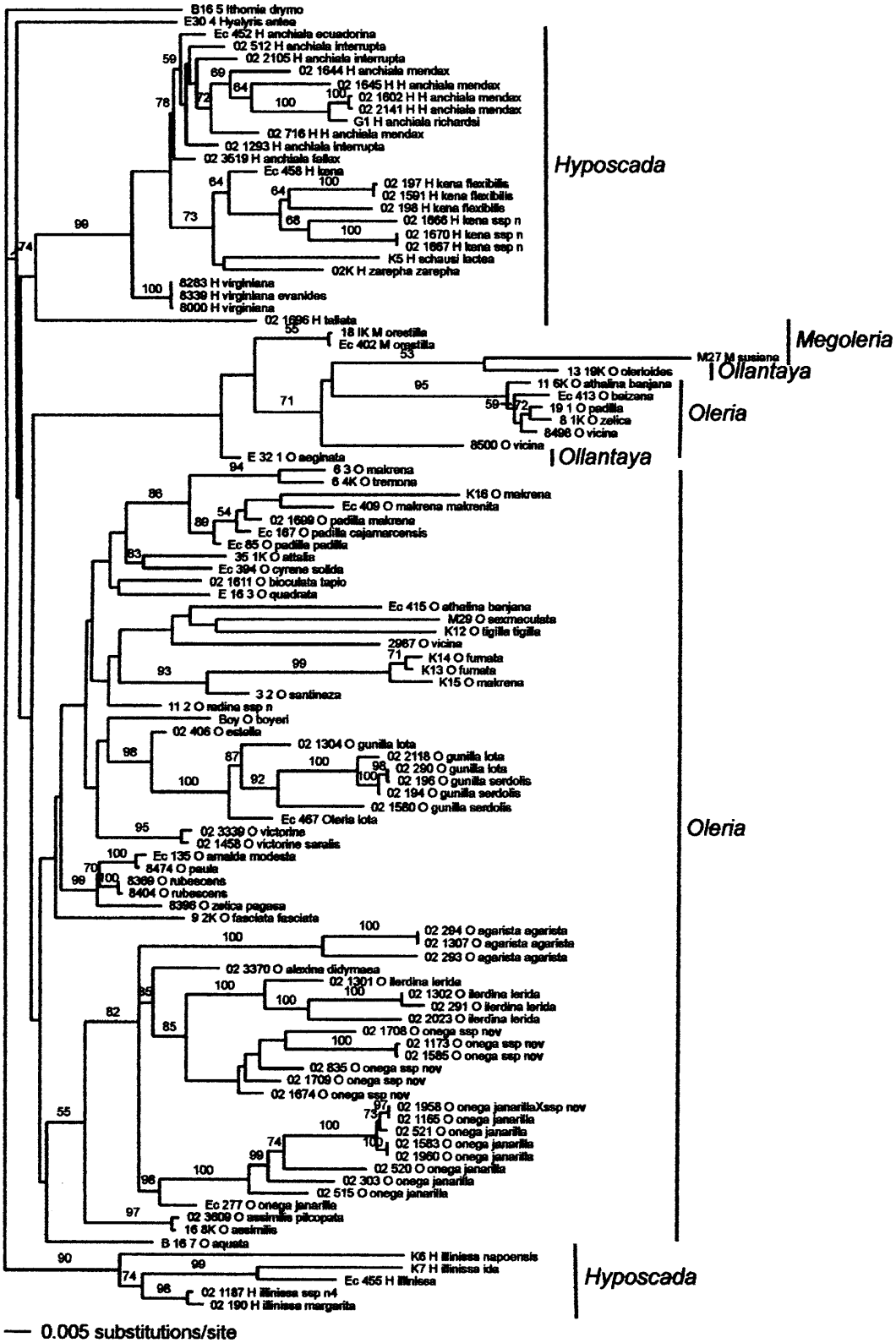


Figure 6. 'Total' data phylogenetic hypothesis of 103 Oleriini specimens, representing the 4 genera (*Hyposcada*, *Megoleria*, *Ollantaya* and *Oleria*), plus 2 outgroup individuals (*Ithomia* and *Hyalyris*). Fragment lengths are as follows: 417 bp *wg*, 1072 bp *Ef1α* and 1618 bp mtDNA. Note that as all available sequence data are included, individual specimens are represented by 1, 2 or all 3 of the *wg*, *Ef1α*, mtDNA gene regions. Tree topology was inferred with neighbour joining.



Based on morphology, four species groups have been determined for the *Oleria* (see Table 3). Unfortunately a sample of *O. aegle*, representing a group on its own, has not yet been obtained for this study. However fieldwork to capture *O. aegle* is planned.

Table 3. *Oleria* species groups based on morphological characters.

	Species group			
	'aegle'	**'amalda'**	**'makrena'**	**'onega'**
***Oleria* species**	*O. aegle*	*O. amalda*	*O. athalina*	*O. agarista*
		O. estella	*O. attalia*	*O. alexina*
		O. gunilla	*O. baizana*	*O. antaxis*
		O. paula	*O. bioculata*	*O. aquata*
		O. rubescens	*O. boyeri*	*O. assimilis*
		O. zelica	*O. cyrene*	*O. astrea*
			O. deronda	*O. didymaea*
			O. derondina	*O. enania*
			O. fasciata	*O. flora*
			O. fumata	*O. ilerdina*
			O. makrena	*O. onega*
			O. padilla	*O. quintina*
			O. phenomoe	*O. sexmaculata*
			O. quadrata	*O. similigena*
			O. radina	*O. synnova*
			O. santineza	*O. thiemei*
			O. tremona	*O. tigilla*
			O. vicina	*O. watkinsi*
			O. victorine	*O. adora*
			O. zea	
			O. bifurcata	
			O. chimera	

The phylogeny determined using *Ef1 α* data (Figure 3) largely supports the division of *Oleria* into the morphologically based species groups, with the 'onega' group sister to the 'amalda' and 'makrena' species. Although the inclusion of *O. aquata* in the 'onega' clade was not supported by bootstrap replication, all other individuals in the 'onega' clade clustered with strong (96%) bootstrap support. The 'amalda' and 'makrena' groups were not well differentiated in molecular analyses; although thirteen 'makrena' species clustered together, *O. boyeri and O. fasciata* (both in the 'makrena' group) clustered with members of the 'amalda' species group. Whilst some bootstrap values within the 'amalda-makrena' grouping were high (e.g. 99% to *O. amalda*, *O. paula*, and *O. rubescens*), branch lengths of the deepest branches were very short and did not attain significant bootstrap support. This is probably in part

attributable to low differentiation between sequences, for example, *O. amalda* (Ec 135) and *O. makrena* (Ec-409) differ at less than 1% of sites in the *Ef1α* region. The slow evolution at this region might also explain why the 'amalda' and 'makrena' groups are not distinguished by *Ef1α*.

The phylogeny obtained using *wg* data (Figure 2) also broadly supports the species groups. The 'onega' + 'amalda' groups cluster with 57% bootstrap support. Except for *O. aquata*, all 'onega' species cluster together with very strong bootstrap support (99%). *O. aquata* assumes a weakly supported position sister to a cluster which contains all 'amalda' species, but the 'amalda' group on their own only attained moderate (55%) bootstrap support. Two 'makrena' species (*O. athalina* and *O. fasciata*) appear to be sisters to the 'onega + amalda' grouping, forming a sister cluster to all other 'makrena' species. The mitochondrial phylogeny (Figure 4) again revealed a distinct 'onega' cluster, with all 'onega' individuals except for *O. aquata* and *O. sexmaculata* clustering with 79% bootstrap support. Members of the paraphyletic 'amalda' and 'makrena' groups cluster together with 53% bootstrap support.

The 'complete' phylogeny (Figure 5) largely mirrors findings from individual gene regions, but often recovers higher support. The 'onega' group clusters together with 98% support (100% for all 'onega' except for *O. aquata*). The 'makrena' and 'amalda' groups cluster together with 61% bootstrap support, with one 'amalda' cluster basal to an 'amalda' + 'makrena' sister pairing. Both 'amalda' clades are supported by 100%, whilst the 'makrena' group receives 67% bootstrap support.

The 'onega' group is repeatedly supported and therefore it is likely that this forms a valid species group. Although *O. aquata* was recovered within the 'onega' group by the 'complete' data set, given that this species is poorly supported by bootstrapping as a member of 'onega' using any individual gene, it would be of great interest to investigate its position in more detail using additional genes and morphological information. The finding that *O. aquata* is distinct from other species in the 'onega' group studied so far is also consistent with its distribution. *O. aquata* is restricted to E. Brazil, S.E. Brazil and N. Venezuela, whilst all other members of the 'onega' group included in the

phylogeny are largely sympatric, found in Colombia, Ecuador, Peru, Bolivia, N. Brazil and N.W. Brazil.

As *O. aquata* and *O. sexmaculata* cluster in the phylogeny, based on mitochondrial sequence data with 88% bootstrap support, and particularly given that sequence data for *O. sexmaculata* were not obtained for either nuclear gene (M29 template DNA was obtained from a dried museum specimen, and is of poor quality), it would be of interest to also obtain further sequence data for this species. Future work should also concentrate on obtaining sequence data to elucidate the phylogenetic positions of *O. similigena*, *O. flora*, *O. astrea* and *O. antaxis*. These would be especially interesting as, like *O. aquata*, they have a more eastern distribution (Venezuela, Guyana, French Guiana and Brazil) than the other 'onega' group species. These four species therefore might be expected to contribute to a more reliable estimate of the phylogenetic position of *O. aquata*, and a more dependable assessment of which species belong in the 'onega' species group.

Conclusions about taxa belonging to, and validity of, the 'amalda' and 'makrena' groups are complicated by possible paraphyly, and the different evolutionary directions recovered by the different gene regions; *Ef1α* suggests the 'makrena' group was derived from the 'amalda' group, whilst *wg* indicates the 'amalda' group might have arisen from the 'makrena' group. What is clear is that there is a very close association between the species belonging to both of these species groups (for example, Figure 5). The molecular data currently available support either grouping all members of 'amalda' and 'makrena' groups into a single species group, or the splitting of the 'amalda' group and also possibly the 'makrena' group.

There is support for the 'amalda' group being split into two groups; the first group would consist of *O. amalda*, *O. paula*, *O. rubescens* and *O. zelica*, and the second group of just *O. estella* and *O. gunilla*. The distribution of these groups overlaps in Colombia and Ecuador, but this split is consistent with biogeography, with the first group assuming more northern distributions (Ecuador, Colombia, Panama, Costa Rica, Nicaragua and further North) and the second group assuming more southern distributions (Colombia, Ecuador, Peru, Bolivia, Brazil). This division is also consistent with the different altitudes

occupied by the butterflies; the first group occupy mid-elevation, trans-Andean sites, whilst *O. estella* and *O. gunilla* are restricted to lowland and Amazonian localities.

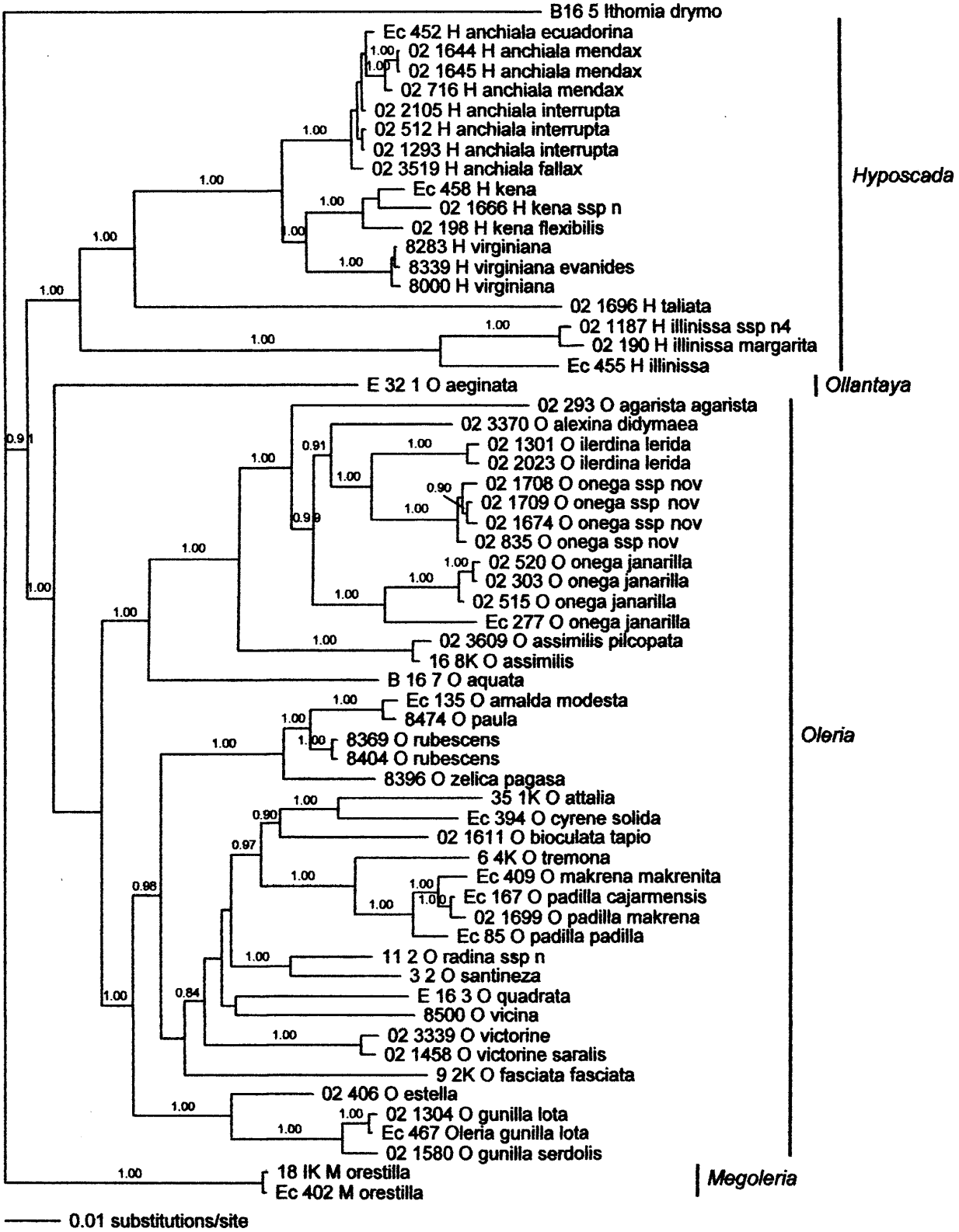
Molecular data from additional individuals of *O. boyeri*, *O. fasciata*, and *O. victorine* are required to determine the validity of a single 'makrena' group. The 'makrena' group typically occupy western locations; Colombia, Ecuador, Peru, Bolivia and Venezuela. However, *O. boyeri* and *O. victorine* have been collected outside this area, in Guyana and S.E. Brazil respectively. In addition, *O. vicina* and *O. zea* are found outside the typical 'makrena' distributional range, being restricted to Central America. *O. vicina* assumes a phylogenetic position within the 'makrena' group, but it remains of interest to investigate whether *O. zea* (for which no sample has been obtained to date) assumes a similar position.

Robust species group delimitation in *Oleria* requires a more thorough coverage of members of this genus. Unfortunately even some key species have not been obtained in this study to date, for example, the above mentioned *O. aegle* and eastern members of the 'onega' group. The missing samples include some of the most difficult Oleriini to obtain, and getting hold of tissue of good enough quality for DNA sequencing is no easy task. However, future work should concentrate on obtaining these samples as each species would be a hugely valuable addition to this work.

Phylogenetic construction method

The phylogenetic hypotheses created for the 'complete' dataset differed depending on the construction method employed. Whilst NJ (Figure 5) and MP (Figure 1) both recovered *Hyposcada* as sister to all other Oleriini, Bayesian analysis (Figure 7) recovered a sister pairing of *Megoleria* and all other Oleriini. This is interesting because, as described above, reports based on both morphological (Willmott, personal correspondence) and higher level molecular (Brower et al. 2005) evidence suggested *Megoleria* to be sister to all other Oleriini. The relationships between some species in the 'makrena' clade (*O. bioculata*, O. *quadrata*, *O. fasciata*, *O. vicina*, *O. victorine*, and *O. radina* + *O. santineza*) also differ significantly depending on the reconstruction method. Elucidating the correct relationships between taxa at all levels within the Oleriini tribe is fundamental to this study. By sequencing the gene regions more

Figure 7. 'Complete' data phylogenetic hypothesis of 60 Oleriini specimens, representing the 4 genera (*Hyposcada*, *Megoleria*, *Ollantaya* and *Oleria*), plus *Ithomia* outgroup. Fragment lengths are as follows: 417 bp *wg*, 1072 bp *Ef1α* and 1618 bp mtDNA. Only those individuals for which data were available for all 3 gene regions are included. Tree topology is consensus of the last 9,000 trees inferred using Bayesian methods. Bayesian posterior probabilities greater than 0.70 are given above branches.



completely, the different reconstruction methods will converge towards the most appropriate phylogeny. This should enable the most accurate phylogenetic hypothesis to be selected and therefore will increase the reliability of subsequent inferences.

Conclusion

Phylogenetic hypotheses were presented for approximately two thirds of the species belonging to the tribe Oleriini, based on sequence data for three gene regions. Although this represents a significant contribution to the molecular systematic revision of the tribe, more accurate phylogenies for the Oleriini could be obtained with further work, for example, by obtaining sequence data for species currently not represented in the study and by including more gene regions. In addition, phylogenetic robustness would be greatly improved by including independent morphological evidence (see Wahlberg et al. 2005).

The recent resurgence of interest in the study of speciation has resulted in the re-exploration of existing hypotheses, as well as the introduction of a number of novel theories, which attempt to account for the origin of species; how, why, where and when? Ultimately, the Oleriini phylogeny will provide a reliable historical framework that can be used to test these, and other, important biological theories. This will provide valuable insights into the biological mechanisms working on a specific tribe, as well as contribute to a better overall understanding of the evolution of Neotropical species.

References

Ackery, P. R. (1988) Hostplants and classification: a review of Nymphalid butterflies. *Biological Journal of the Linnean Society,* **33**: 95-203.

Baker, R. H. & DeSalle, R. (1997) Multiple sources of character information and the phylogeny of Hawaiian *Drosophila*. *Systematic Biology*, **46**: 654-673.

Barkman, T. J. & Simpson, B. B. (2001) Origin of high-elevation *Dendrochilum* species (Orchidaceae) endemic to Mount Kinabalu, Sabah, Malaysia. *Systematic Biology*, **26** (3): 658-669.

Brower, A. V. Z. & DeSalle, R. (1998) Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of *wingless* as a source of characters for phylogenetic inference. *Insect Molecular Biology,* **7**: 1-10.

Brower, A. V. Z., Freitas, A. V. L., Lee, M.-M., Silva Brandão, K. L., Whinnett, A. & Willmott K. R. (2006) Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions. *Systematic Entomology*, **31**: 288-301.

Brown, W. L. (1957) Centrifugal speciation. *Quarterly Review of Biology*, **32**: 247-277.

Brown, K. S. Jr., & Freitas, A. V. L. (1994) Juvenile stages of Ithomiinae: overview and systematics. *Tropical Lepidoptera*, **5**: 9-20.

Cho, S., Mitchell, A., Regier, J. C., Mitter, C., Poole, R. W., Friedlander, T. P. & Zhao, S. (1995) A highly conserved nuclear gene for low-level phylogenetics: elongation factor 1α recovers morphology-based tree for heliothine moths. *Molecular Biology and Evolution*, **12**: 650-656.

Fox, R. M. (1968) *Ithomiidae* (Lepidoptera: Nymphaloidea) of Central America *Trans. Am. Entomol. Soc.* **94**: 155-208.

Fjeldsä, J. (1994) Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodiversity and Conservation*, **3**: 207-226.

Harvey, D. J. (1991) Higher classification of the Nymphalidae. *The Development and Evolution of Butterfly Wing Patterns* (Ed.) H. F. Nijhout, pp. 255-273. Smithsonian Institution Press, Washington D. C.

Huelsenbeck, J. P. & Ronquist F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**: 754-755.

Lamas, G. (2004) Nymphalidae. Ithomiinae, pp. 172-191. In G. Lamas (ed.), Checklist: Part 4A. Hesperioidea - Papilionoidea. In (Ed.) J. B. Heppner, *Atlas of Neotropical Lepidoptera. Volume 5A.* Association for Tropical Lepidoptera, Scientific Publishers, Gainesville.

Motta, P. C. (2003) Phylogenetic relationships of Ithomiinae based on first-instar larvae. In *Butterflies: ecology and evolution taking flight.* (Ed.) C. L. Boggs, W. B. Watt and P. R. Ehrlich), pp. 409-429. University of Chicago Press, Chicago.

Roy, M. S. (1997) Recent diversification in African greenbuls (Pycnonotidae: Andropadus) supports a montane speciation model. *Proceedings of the Royal Society of London. Series B*, **263:** (1386): 1337-1344.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. (1994) Evolution, weighting, and phylogenetic usefulness of mitochondrial genes with a compilation of conserved PCR primers. *Annals of the Entomological Society of America*, **87**: 651-701.

Sorenson, M. D. (1999) TreeRot, version 2. Boston University, Boston, Massachusetts.

Swofford, D. L. (2000) *PAUP**: Phylogenetic Analysis Using Parsimony (* and Other Methods). Sinauer Associates, Sunderland, Massachusetts.

Wahlberg, N., Braby, M. F., Brower A. V. Z., de Jong, R., Lee, M.-M., Nyline, S., Pierce, N. E., Sperling, F. A. H., Vila, R., Warren, A. & Zakharov, E. (2005) Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers. *Proceedings of the Royal Society, Series B*, **272**: 1577-1586.

Whinnett, A., Brower, A. V. Z., Lee, M.-M., Willmott, K. R. & Mallet, J. (2005) The phylogenetic utility of *Tektin*, a novel region for inferring systematic relationships amongst Lepidoptera. *Annals of the Entomological Society of America*, **98**: 873-886.

CHAPTER SEVEN

CONCLUSIONS

Phylogeography

The term 'phylogeography' was introduced in 1987 by Avise and colleagues to describe the section of biogeography which is specifically 'concerned with the principles and processes governing the geographic distributions of genealogical lineages' (Avise 2000). In the final section of the authorative book entitled 'Phylogeography', Avise (2000) highlighted three directions for future work in this field: 1) the use of nuclear DNA, 2) an increase in the integration of molecular data with information of profound relevance to phylogeographic patterns, and 3) the application of 'comparative phylogeography on a regional scale, using multiple co-distributed species'. The research presented here began just one year after Avise's (2000) book was published. In an attempt to undertake research at the forefront of phylogeography, I aimed to incorporate each of Avise's recommendations into my study of ithomiine butterflies.

1) The use of nuclear DNA

Due to difficulties with primer design, such as large intron insertions and a lack of available, alignable sequences, at the start of this thesis few nuclear loci had been used in molecular systematics studies of Lepidoptera. In this thesis I further developed the nuclear locus *triosephosphate isomerase* (*Tpi*) (Chapter 3), from the exon 3 to 4 region which had previously been used by Beltrán and colleagues (2002) in their studies of *Heliconius* butterflies, so that the region between exons 1 and 5 could be amplified. I also developed a new nuclear region, *Tektin* (Chapter 4), for research into molecular evolution, phylogeny and phylogeography. These add to the primers available for future studies on Lepidoptera. Indeed, *Tektin* has already proved useful in other species level studies, for example in the genera *Ithomia* (Mallarino et al. 2005), *Melinaea* and *Hypothyris* (Zimmermann et al. pers. comm.). In addition, I sequenced two tried and tested nuclear regions for the phylogeny of Oleriini (Chapter 6), which forms

the basis of the historical framework which will be used in future phylogeographic analyses of this tribe.

2) An increase in the integration of molecular data with information of profound relevance to phylogeographic patterns.

The most biologically relevant phylogeographic conclusions are drawn in those studies which interpret the molecular data in conjunction with additional knowledge about the study taxa, such as specifics about the distributional patterns of the taxa, or aspects of their natural history. In this thesis, I included such knowledge both in the study design and also when interpreting molecular results.

For example, in the work on the suture zone in N.E. Peru (Chapters 2 & 3), knowledge of which collection localities belonged to each of the Ucayali and Huallaga zones of endemism was integral to the study design. Taxon pairs were not just selected from those belonging to any two different collection localities, but were specifically selected from those belonging to the two different centres of endemism. For example, three individuals of *Hyposcada anchiala* were selected from the five collection localities in the Huallaga zone. An additional three individuals of *Hyposcada anchiala* were selected from the three collection localities in the Ucayali zone. By selecting taxa from both zones of endemism, or postulated 'refuges', I was able to calculate divergences between taxa belonging to different 'refuges' which therefore enabled a test of the Pleistocene refuge theory, rather than just to identify more general patterns.

I also used additional information (wing pattern, geographical distribution, and altitudinal zone) to interpret the molecular phylogeny of *Hyposcada anchiala* (Chapter 5). In the future, detailed ecological information about host-plant and other habitat requirements will be incorporated in work of this kind (e.g. Willmott and Mallet 2004).

3) The application of 'comparative phylogeography on a regional scale, using multiple co-distributed species'

Phylogeography gains power as a method for finding general patterns when a multi-species approach is used. Such work attempts to investigate whether co-distributed taxa share concordant signatures that indicate a shared history. An important type of information for such replicated, comparative studies can come from 'suture zones' (zones where multiple pairs of co-distributed taxa interact in contact zones), such as studied here. In this thesis, I investigated 172 ithomiines, which had been collected from a study area approximately 90 x 55 km, for mitochondrial DNA, (Chapter 2) and 95 for *Tpi* (Chapter 3), as part of the most comprehensive, regional phylogeographic study performed on Neotropical taxa and, I believe of any suture zone, to date.

I found that molecular divergence, and therefore inferred divergence times, between the ithomiine butterflies across the N.E. Peru suture zone differed remarkably. I therefore infer that divergences in the ithomiines studied here were not driven by simultaneous vicariance. This finding adds to the growing body of evidence that the climatic fluctuations of the Pleistocene impacted Neotropical and temperate taxa in very different ways.

In addition, this work addresses a gap in genetic divergence research which was identified nearly 10 years ago, but still largely holds true today: 'There are as yet few studies which allow a comparison of DNA divergence across racial, specific and generic taxonomic levels in the same group' (Hewitt 1996).

Finally

There has been a great deal of speculation about which factors drove the generation of the unrivalled terrestrial biodiversity of the Neotropical forests, but a consensus has not yet been reached. I believe that this thesis makes a valuable contribution to the current understanding of Neotropical diversification. A greater understanding of which processes contributed to the Neotropical biodiversity will enable more effective conservation strategies to be developed. Conservation strategies are desperately needed to protect the Neotropical forest

biome, which is being destroyed at such an alarming rate. And in particular, to protect those forest areas which contain the endemic taxa, amongst them the ithomiines, that we are in imminent danger of losing forever.

Phylogeography is undoubtedly a powerful approach which enables informed comments to be made about the events which shaped the patterns of genetic variation that we can detect today. I have successfully incorporated all three of the directions identified for the future of phylogeography by Avise in 2000. I argue that the results presented here are at the forefront of this dynamic research field.

References

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. & Saunders, N. C. (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu Rev. Ecol Syst.*, **18**: 489-522.

Avise, J. C. (2000) Phylogeography. The history and formation of species. Harvard University Press. London, England.

Beltrán, M. S., Jiggins, C. D., Bull, V., Linares, M., Mallet, J., McMillan, W. O. & Bermingham, E. (2002) Phylogenetic discordance at the species boundary: comparative gene genealogies among rapidly radiating *Heliconius* butterflies. *Mol. Biol. Evol.*, **19**: 2176-2190.

Hewitt, G. M. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc.*, **58**: 247-276.

Mallarino, R., Bermingham, E., Willmott, K. R., Whinnett, A. & Jiggins, C. D. (2005) Molecular systematics of the butterfly genus *Ithomia* (Lepidoptera: Ithomiinae): a composite phylogenetic hypothesis based on seven genes. *Mol. Phylogenet. Evol.*, **34**: 625-644.

Willmott, K. R. & Mallet, J. (2004) Correlations between adult mimicry and larval hostplants in ithomiine butterflies. *Proc. Roy. Soc. Lond. B (Suppl.)*, **271**: S266-S269.

APPENDIX ONE

MOLECULAR SYSTEMATICS OF THE BUTTERFLY GENUS *ITHOMIA* (LEPIDOPTERA: ITHOMIINAE): A COMPOSITE PHYLOGENETIC HYPOTHESIS BASED ON SEVEN GENES